

```

1   CGGGCGCGGC GGC GGCGGCG GTGACAGCG CGCCCGCGCC TCCCCGCGCG
51  TAGGTGTGCG GCGCGCTCCT GGCGAGGACG GAGCGAGCAG ATCTCGCGTG
101 CGCTCGCCGC CCGGCGCAGC CCAGCCCGGC CCCC GCCTGG CGCCGCGAGC
151 CGAGGTGTCT CCCGCGCCCG CGCCCGTGTC GCCGCGGTGC CCGCGAGCGG
201 GAGCCGGAGT CGCCGCGGCC CGAGCGCAGC CGAGCGCACG CCGAGCCCGT
251 CCGCCGCGCG CATGGCCACC ACGGTGACCT GCACCCGCTT CACCGACGAG
301 TACCAGCTCT ACGAGGATAT TGGCAAGGGG GCTTTCTCTG TGGTCCGACG
351 CTGTGTCAAG CTCTGCACCG GCCATGAGTA TGCAGCCAAG ATCATCAACA
401 CCAAGAAAGCT GTCAGCCAGA GATCACCAGA AGCTGGAGAG AGAGGCTCGG
451 ATCTGCCGCC TTCTGAAGCA TTCCAACATC GTGCGTCTCC ACGACAGCAT
501 CTCCGAGGAG GGCTTCCACT ACCTGGTCTT CGATCTGGTC ACTGGTGGGG
551 AGCTCTTTGA AGACATTGTG GCGAGAGAGT ACTACAGCGA GGCTGATGCC
601 AGTCACTGTA TCCAGCAGAT CCTGGAGGCC GTTCTCCATT GTCACCAAAT
651 GGGGGTCTGC CACAGAGACC TCAAGCCGGA GAACCTGCTT CTGGCCAGCA
701 AGTGCAAAGG GGCTGCAGTG AAGCTGGCAG ACTTCGGCCT AGTATCGAG
751 GTGCAGGGGG ACCAGCAGGC ATGGTTTGGT TTCGCTGGCA CACCAGGCTA
801 CCTGTCCCTT GAGGTCTTTC GCAAAGAGGC GTATGGCAAG CCTGTGGACA
851 TCTGGGTCATG TGGGGTGATC CTGTACATCC TGCTCGTGGG CTACCCACCC
901 TTCTGGGACG AGGACCAGCA CAAGCTGTAC CAGCAGATCA AGGCTGGTGC
951 CTATGACTTC CCGTCCCTTG AGTGGGACAC CGTCACTCCT GAAGCCAAAA
1001 ACCTCATCAA CCAGATGCTG ACCATCAACC CTGCCAAGCG CATCACAGCC
1051 CATGAGGCCC TGAAGCACCC GTGGGTCTGC CAACGCTCCA CGGTAGCATC
1101 CATGATGTCAC AGACAGGAGA CTGTGGAGTG TCTGAAAAAG TTCAATGCCA
1151 GGAGAAAGCT CAAGGGAGCC ATCCTCACCA CCATGCTGGC CACACGGAAT
1201 TTCTCAGTGG GCAGACAGAC CACCGCTCCG GCCACAATGT CCACCGCGGC
1251 CTCCGGCACC ACCATGGGGC TGGTGGAAAC AGCCAAGAGT TACTCAACA
1301 AGAAAGCAGA TGGAGTCAAG CCCCAGACGA ATAGACCAA AAACAGTGCA
1351 GCCGCCACCA GCCCCAAAGG GACGCTTCCT CCTGCCGCCC TGGAGCCTCA
1401 AACCACCGTC ATCCATAACC CAGTGGACGG GATTAAGGAG TCTTCTGACA
1451 GTGCCAATAC CACCATAGAG GATGAAGACG CTAAGCCCG GAAGCAGGAG
1501 ATCATTAAGA CCACGGAGCA GCTCATCGAG GCCGTCAACA ACGGTGACTT
1551 TGAGGCCTAC GCATTCTACT TCGAGAACCT GCTGGCCAAG AACAGCAAGC
1601 CGATCCACAC GACCATCCTG AACCACACG TGCACGTCAT TGAGAGGAT
1651 GCCGCCTGCA TCGCTTACAT CCGGCTCACG CAGTACATTG ACGGGCAGGG
1701 CCGGCCCCGC ACCAGCCAGT CTGAGGAGAC CCGCGTGTGG CACCGCCGCG
1751 ACGGCAAGTG GCAGAACGTG CACTTCCACT GCTCGGGCGC GCCTGTGGCC
1801 CCGCTGCAGT GAAGCCAAGG GAGGGGCACA GAATGGGGAA CAGGACACAG
1851 GATCCTAAAC TCCAAGGGGA CTGTCCACCG ATGAACACTC AGAGTGGACA
1901 CCATCTTCCG TCCACGCTGT GCCCAGGACA GCTGTCCCA TCCATGAACA
1951 CAGGGTAAAC ATCTGCCGGG CTCCGCACCA GTGGCTCCCT GGGCCATGGG
2001 ACAGCGGCAG GGCTCACCAC GGACAGCACG TGGCCCAGCA GCCGGCCACC
2051 CTGGCGTCCT GGGGCTCCT CCCCTCCTCT CCCTCTCACC TTGTACCTC
2101 CACGGAGCTG CTTGTCTGGG ATAATTTGGG GATTTT TTTT TCTGGGGGAT
2151 AATTCTTTTG CATGACCCCT AAAGAGCAAG CCACACCGGT CTGCTAGCTA
2201 GGTGTCCGCG GTGTGGTG (SEQ ID NO:1)

```

FEATURES:

```

5'UTR:      1-261
Start Codon: 262
Stop Codon:  1810
3'UTR:      1813

```

Homologous proteins:

Top 10 BLAST Hits

			Score	E
CRA 18000005245285	/altid=gi 5326757 /def=gb AAD42035.1 AF07880...		1047	0.0
CRA 18000005199792	/altid=gi 10835006 /def=ref NP_001211.1 cal...		1044	0.0
CRA 18000004938668	/altid=gi 6671660 /def=ref NP_031621.1 calc...		1039	0.0
CRA 18000004937301	/altid=gi 11120682 /def=ref NP_068507.1 Ca+...		1038	0.0
CRA 18000005245287	/altid=gi 5326762 /def=gb AAD42037.1 AF08192...		1001	0.0
CRA 18000005171302	/altid=gi 3668373 /def=gb AAC79460.1 (AF085...		999	0.0
CRA 1000737074531	/altid=gi 6688228 /def=emb CAB65122.1 (AJ252...		986	0.0
CRA 18000005245288	/altid=gi 5326764 /def=gb AAD42038.1 AF08341...		986	0.0
CRA 18000004964693	/altid=gi 466360 /def=gb AAA81938.1 (U06636...		982	0.0
CRA 18000005199791	/altid=gi 4139268 /def=gb AAD03743.1 (AF112...		982	0.0

BLAST dbEST hits:

	Score	E
gi 12801212 /dataset=dbest /taxon=960...	1675	0.0
gi 12868201 /dataset=dbest /taxon=960...	1453	0.0
gi 2053138 /dataset=dbest /taxon=9606 ...	1247	0.0
gi 10213950 /dataset=dbest /taxon=96...	1243	0.0
gi 9324431 /dataset=dbest /taxon=960...	1233	0.0
gi 12921378 /dataset=dbest /taxon=960...	910	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|12801212 Fetal brain

gi|12868201 Fetal brain

gi|2053138 Testis

gi|10213950 Lung small cell carcinoma

gi|9324431 uterus endometrium adenocarcinoma cell libe

gi|12921378 Fetal brain

Tissue expression from PCR-based tissue screening panels:

hippocampus

```

1 MATTVTCTRF TDEYQLYEDI GKGAFSVVRR CVKLCTGHEY AAKIINTKKL
51 SARDHQKLER EARICRLKH SNIVRLHDSI SEEGFHYLVF DLVTGGELFE
101 DIVAREYYSE ADASHCIQQI LEAVLHCHQM GVVHRDLKPE NLLLASKCKG
151 AAVKLADFGI AIEVQGDQQA WFGFAGTPGY LSPEVLRKEA YGKPVDIWAC
201 GVILYILLVG YPPFWDEDQH KLYQQIKAGA YDFPSPEWDT VTPEAKNLIN
251 QMLTINPAKR ITAHEALKHP WVCQRSTVAS MMHRQETVEC LKKFNARRKL
301 KGAILTTMLA TRNFSVGRQT TAPATMSTAA SGTTMGLVEQ AKSLNKKAD
351 GVKPQTNSTK NSAAATSPKG TLPPAALEPQ TTVIHNPVDG IKESSDSANT
401 TIEDEDAKAR KQEIIKTTEQ LIEAVNNGDF EAYAFYFENL LAKNSKPIHT
451 TILNPHVHVI GEDAACIAYI RLTOYIDGQG RPRTSQSEET RVWHRDQKW
501 QNVHFHCSGA PVAPLQ (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 3

```

1 313-316 NFSV
2 357-360 NSTK
3 399-402 NTTI

```

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 2

```

1 48-51 KKLS
2 259-262 KRIT

```

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 4

```

1 47-49 TKK
2 51-53 SAR
3 358-360 STK
4 367-369 SPK

```

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 9

```

1 36-39 TGHE
2 51-54 SARD
3 79-82 SISE
4 94-97 TGGE
5 109-112 SEAD
6 262-265 TAHE
7 400-403 TTIE
8 401-404 TIED
9 485-488 SQSE

```

[5] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

9-17 RFTDEYQLY

[6] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 3

1	302-307	GAILTT
2	332-337	GTTMGL
3	390-395	GIKESS

[7] PDOC00100 PS00107 PROTEIN_KINASE_ATP
Protein kinases ATP-binding region signature

20-43 IGKGAFSVVRRVCVKLCTGHEYAAK

[8] PDOC00100 PS00108 PROTEIN_KINASE_ST
Serine/Threonine protein kinases active-site signature

132-144 VVHRDLKPENLLL

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	195	215	1.665	Certain
2	319	339	1.301	Certain

BLAST Alignment to Top Hit:

>CRA|18000005245285 /altid=gi|5326757 /def=gb|AAD42035.1|AF078803_1
 (AF078803) calcium/calmodulin-dependent protein kinase II
 beta subunit; CAM2 [Homo sapiens] /org=Homo sapiens
 /taxon=9606 /dataset=nraa /length=542
 Length = 542

Score = 1047 bits (2678), Expect = 0.0
 Identities = 516/542 (95%), Positives = 516/542 (95%), Gaps = 26/542 (4%)
 Frame = +1

Query: 1 MATTVTCTRTFTDEYQLYEDIGKGAFSVVRRCKLCTGHEYAAKIINTKKLSARDHQKLER 180
 MATTVTCTRTFTDEYQLYEDIGKGAFSVVRRCKLCTGHEYAAKIINTKKLSARDHQKLER
 Sbjct: 1 MATTVTCTRTFTDEYQLYEDIGKGAFSVVRRCKLCTGHEYAAKIINTKKLSARDHQKLER 60

Query: 181 EARICRLKHSNIVRLHDSISEEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIQQI 360
 EARICRLKHSNIVRLHDSISEEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIQQI
 Sbjct: 61 EARICRLKHSNIVRLHDSISEEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIQQI 120

Query: 361 LEAVLHCHQMGGVVRDLKPENLLLASKCKGAAVKLADFLAIEVQGDQQAQWFGFAGTPGY 540
 LEAVLHCHQMGGVVRDLKPENLLLASKCKGAAVKLADFLAIEVQGDQQAQWFGFAGTPGY
 Sbjct: 121 LEAVLHCHQMGGVVRDLKPENLLLASKCKGAAVKLADFLAIEVQGDQQAQWFGFAGTPGY 180

Query: 541 LSPEVLRKEAYGKPVDIWACGVILYILLVGYPFFWDEDQHKLYQQIKAGAYDFPSPWDT 720
 LSPEVLRKEAYGKPVDIWACGVILYILLVGYPFFWDEDQHKLYQQIKAGAYDFPSPWDT
 Sbjct: 181 LSPEVLRKEAYGKPVDIWACGVILYILLVGYPFFWDEDQHKLYQQIKAGAYDFPSPWDT 240

Query: 721 VTPEAKNLINQMLTINPAKRITAHEALKHPWVCQRSTVASMMHRQETVECLKKFNARRKL 900
 VTPEAKNLINQMLTINPAKRITAHEALKHPWVCQRSTVASMMHRQETVECLKKFNARRKL
 Sbjct: 241 VTPEAKNLINQMLTINPAKRITAHEALKHPWVCQRSTVASMMHRQETVECLKKFNARRKL 300

Query: 901 KGAILTTMLATRNFSVGRQTTAPATMSTAASGTTMGLVEQAKSLNKKADGVKPQTNSTK 1080
 KGAILTTMLATRNFSVGRQTTAPATMSTAASGTTMGLVEQAKSLNKKADGVKPQTNSTK
 Sbjct: 301 KGAILTTMLATRNFSVGRQTTAPATMSTAASGTTMGLVEQAKSLNKKADGVKPQTNSTK 360

Query: 1081 NSAAATSPKGTLPAALEPQTTVIHNPVDGIKESSDSANTTIEDEDAKARKQEIIKTTEQ 1260
 NSAAATSPKGTLPAALEPQTTVIHNPVDGIKESSDSANTTIEDEDAKARKQEIIKTTEQ
 Sbjct: 361 NSAAATSPKGTLPAALEPQTTVIHNPVDGIKESSDSANTTIEDEDAKARKQEIIKTTEQ 420

Query: 1261 LIEAVNNGDFEAYA-----FYFENLLAKNSKPIHTTILN 1362
 LIEAVNNGDFEAYA FYFENLLAKNSKPIHTTILN
 Sbjct: 421 LIEAVNNGDFEAYAKICDPGLTSFEPEALGNLVEGMDFHRFYFENLLAKNSKPIHTTILN 480

Query: 1363 PHVHVIGEDAACIAYIRLTQYIDGQGRPRTSQSEETRVWHRRDGKWQNVHFHCSGAPVAP 1542
 PHVHVIGEDAACIAYIRLTQYIDGQGRPRTSQSEETRVWHRRDGKWQNVHFHCSGAPVAP
 Sbjct: 481 PHVHVIGEDAACIAYIRLTQYIDGQGRPRTSQSEETRVWHRRDGKWQNVHFHCSGAPVAP 540

Query: 1543 LQ 1548
 LQ
 Sbjct: 541 LQ 542 (SEQ ID NO:4)

FIGURE 2, page 3 of 4

Hammer search results (Pfam):

Model	Description	Score	E-value	N
PF00069	Eukaryotic protein kinase domain	306.2	3.9e-88	1
CE00022	CE00022 MAGUK_subfamily_d	293.8	1.3e-86	1
CE00359	E00359 bone_morphogenetic_protein_receptor	15.0	0.0015	1
CE00031	CE00031 VEGFR	0.9	2.1	1
CE00287	CE00287 PTK_Eph_orphan_receptor	-65.4	0.00046	1
CE00292	CE00292 PTK_membrane_span	-77.0	0.00018	1
CE00291	CE00291 PTK_fgf_receptor	-93.1	0.0021	1
CE00286	E00286 PTK_EGF_receptor	-132.2	0.0059	1
CE00290	CE00290 PTK_Trk_family	-161.3	0.00033	1
CE00016	CE00016 GSK_glycogen_synthase_kinase	-196.7	9.2e-06	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
CE00359	1/1	132	186 ..	272	327 ..	15.0	0.0015
CE00031	1/1	133	205 ..	1068	1139 ..	0.9	2.1
CE00286	1/1	14	252 ..	1	263 []	-132.2	0.0059
CE00290	1/1	15	253 ..	1	282 []	-161.3	0.00033
CE00291	1/1	14	267 ..	1	285 []	-93.1	0.0021
CE00292	1/1	14	267 ..	1	288 []	-77.0	0.00018
CE00287	1/1	14	270 ..	1	260 []	-65.4	0.00046
PF00069	1/1	14	272 ..	1	278 []	306.2	3.9e-88
CE00022	1/1	10	305 ..	13	316 ..	293.8	1.3e-86
CE00016	1/1	1	343 [.	1	433 []	-196.7	9.2e-06

1 GAGCTGCTGT GTCTCTGTCC CCAGGGGCGAG AGGGGCTGTG GGGTTGCAGG
 51 CTAGCGTCT GGGACTCTGG GGTGAAGGCT CAGCCATGCC CTGCAGACAC
 101 CATGGGGCAG GGCTCAGACC TGTGCACCTG TCTCTTGCAA ACCACTGTTT
 151 TCTCTGTTTT GTAACCCCC ACCCAACCCC ACATAACACC TCTGGGTTTA
 201 AACACATGC ACCCTTGTGC CGGTACCTC CCTGCAGCCG GAGAACCTGC
 251 TTCTGGCCAG CAAGTGCAAA GGGGCTGCAG TGAAGCTGGC AGACTTCGGC
 301 CTAGCTATCG AGGTGCAGGG GGACCAGCAG GCATGGTTTG GTGAGTGCCA
 351 GGGGCAGGGT GTGTTGGCTG GCAGTTGGCA GGGCAGGAGG TGATGCTGAC
 401 AGCCCTTGT GGCCTCTTCC CCTCTCTCTA GGTTTCGCTG GCACACCAGG
 451 CTACCTGTCC CCTGAGTTC TTCGCAAAGA GCGTATGGC AAGCCTGTGG
 501 ACATCTGGGC ATGTGGTGAG GCCTGGCCTG AGTTGGTGCG GGGCAGGGCC
 551 TCGGGTGTTT CAGGACTTCC CACCTACATC CTGGAGTGTG CAGTGGCCAG
 601 CACGTCTTGC TCTCATCTGG GTTTATCTGT GTCAGACCTG CCCTTGAGCT
 651 GCCCTGGCAG GGGTCTGCCC ACACAGCCAA GAGCCCCCTT TCCACCCAGA
 701 TTAGAATTGC TCACATGAAC CTGGCGCACC CCAGTGCTCG CCTGCGCTCA
 751 GCAGAGGTCT GTCCAGAAAG TGTGGTGGGT GGATGGGAGT GGAGAAGAGA
 801 GGTCAGGGGC TGTGGGCCA TGGGCAGGGC CACCTCCTTG GGTAGGGGTC
 851 TCCTCCCA CAAGGTGGGA GCAGCAGAGG GGCTTGACAT CACCCTCATC
 901 CCTGTGATAG TGTGGGTGTG GGCAGAGGT CAGGGGGCCG GCTGTGCCCT
 951 TCTACCCAG TGTCTGTGC ACAGGTGGGG GCAAAGGAAT GCTGAGGACC
 1001 CCAATGCCCT CCCAGGGCCA CAGGAGCTAG GCAGTGAGGG TGCAGGGCAT
 1051 GGGCTTCATG GACGGTGGCA CCCTGCAAGT GGCTGCGGTG CTCACAGGCC
 1101 CCATCCGCGAG GGGTGATCCT GTACATCCTG CTCGTGGGCT ACCCACCCTT
 1151 CTGGGACGAG GACCAGCACA AGCTGTACCA GCAGATCAAG GCTGGTGCCT
 1201 ATGACGTGAG TGCACCAGCC CCTCTCTGAT GAGCTCCCTT CCTCCAGGTG
 1251 TGGCCGGGTG AGGGCAGCGT GGGAGAGGC TAGGAGTGGG GTGAAGCCAC
 1301 CTGTGGCCAG GTCCTGGGTC CTGCTCTCCC AGATTCTGTG CTGGAGATGA
 1351 AGCCCTTGG AGAATTCTTG CCCCTGCCTG AGAGGGAGCT TCAGGCCCGG
 1401 CCGGGCGCTG GTTCTCTTCT GCAGTTCCCG TCCCTGAGT GGGACACCGT
 1451 CACTCCTGAA GCCAAAACC TCATCAACCA GATGCTGACC ATCAACCCTG
 1501 CCAAGCGCAT CACAGCCCAT GAGGCCCTGA AGCACCCTG GGTCTGCGTG
 1551 AGTCGCCCTT GGTGCCCATG GTGGGGAGGG GGCTCCTGGT GGAGATGGCC
 1601 TCAGACCACT CCCCTGGCAA GGACCCCAAG AGGGTCCTGT TCCTGACATC
 1651 CAAGAGCTCC CTTGGGTCCC CTGGGTGCTC CTTGTGGCTT CTGGCTTGGG
 1701 ACATACCAGC ACGTTTGTA GGCCTGGGGC TTGGAAGGCA TTAGAGGGTA
 1751 GAGGTGATCC CTTCCCTCCA ACTGCAGTCC TGTCTGTGAG GGGCAGAGTG
 1801 GACGAGGCAA GGGAGAGACG AGTCTTGAAG TCCCAGGCGG GTGGGGACAG
 1851 ACAACCCTTG CCGCAATGGT GGCCGGTGGC TCTTGGAAG TGGGGACCCC
 1901 AGGGTGCCAC AAGCCTTGCC ACCCTGGCCT CTCCTCTGTG CCTCGGGCTC
 1951 GGCTGCCATA TGACCACCCA TTTCCCCACA GCAACGCTCC ACGGTAGCAT
 2001 CCATGATGCA CAGACAGGAG ACTGTGGAGT GTCTGAAAAA GTTCAATGCC
 2051 AGGAGAAAGC TCAAGGTGAG GCCCTGGCCC CTAGTCCCAG GCACGGCCAT
 2101 GCTTCTCTGT GTCCCCTTGG GCTGGAGCAG GGGGGCCTTG GGGGGTCTGG
 2151 GCAGACCTAG GGGTTACTGC TGCCCCAAG ACTGACTGTT AGCAAGTCCC
 2201 AGACTGGATG CATCAGGTGA ACTCAGGCCA GCTTGGGAAT GAGTCCAGAG
 2251 GGGCCCTGGG CCAGGTGTGG CTCCTCCTAG TTGTCTGTGC CACCTCCTAG
 2301 CAGCCCTTGG AGGAGCTGTC CTGAAGCGCT CGCTGTGGGC TCCTACCCG
 2351 GGCTCTGCAG GCAGCACTCA CCCTCTGGCA GTCACACTGT TTAGTACAAG
 2401 CAAGTCCGAA GCTTCCGGCT CAGACAGGTT TGGTAAGGAG AGCAGAGCCA
 2451 CACACACTGG TCTTGGGTGG GCTGGGGGAG TTCTGGGAGG GAGGTGGGTC
 2501 CCAGTAGGGT ATCCAACCTG CCTGCTTTGG TCAGGGCTGG CTCCGGTGAC
 2551 CGCACACTGG CAGTCCCTCT ACTTGTGGGT TCCGGGATGG GGACTTGTG
 2601 CCTGACTGCC CTCTGCTGGT CTCTGAGCAG TTCTCCCCGG AAGCCCCAGG
 2651 ACTGTTGCCC TGTCTGAGCC TGTGAGGAAA AGAAGGGGCT GTCAGGGAGC
 2701 TGGACCCAG AGGAGCTGCC GTGGTGACCA GCTGTTCTGG TGACCCCTGA
 2751 GGCTTGAGGG GTCTTGAAGC AGCTAGAAGC TGTAAGTGGT CAACAGGTTT
 2801 AGGCCAGGG TGTGTGTAGT TCTGGAAATA GGTGATCTGT CTCAGTCCGG
 2851 CTGTGGCTT CCTGGAGCTC TTGCCTCTCT GGAAGGCTGA GGTGATGCA
 2901 GCCTCATGAC AATGAGGCTG AGCATCTGGG CAGGAGGACA GGGGTCTTAT
 2951 CCTGGCCAGA AGCCAGCAGG GAACACTGAT GGGATAGCCC CGGTTTTATC
 3001 TGTGTCTCT CCCAGGGAGC CATCCTCACC ACCATGCTGG CCACACGGAA
 3051 TTTCTCAGT GAGCCTTTCT TCTCCAGGGA GACAGGCGCT GCCCCCTCCC
 3101 TGCTGGCCCA CGCAGGAGAG CGCCTCCTTC CTCACCAGCC TCTCCACTCC

FIGURE 3, page 1 of 16

3151 TCCTCTGCGG CAGGCCTGCC CTCGGCGTCT GCCCTCAGCT CTGAGACCCA
 3201 CTGCCCCACCT GGCCCCGCTG GGCTCCCACC TTGGGTGATA CCACAGGGTC
 3251 CAGCCCCCCC AGGCCATCAC CTTCTGTCTG GGTCTGTGTC CCTCCACCCC
 3301 CTGAACACGA GCGTCTGTGC TGCCCCACTG GGGCTCACAG CATCGTGTGT
 3351 GTCTGTCCAG GCGTTTGTGC GGCATCTATG TGGCTCCTT GTCATTTTGA
 3401 GTGCTCTGAA CATTGTGTTT TGTGCGGGAG GTGGGCAGAA GGGATGCGGG
 3451 GTGATGCGGG AGGCTCGGGG GCCTCCTTCC AAGTTCTGGA TGAGCTGCAG
 3501 CCTCCTGTCC CGGCTGCTCA GGGTGGGTGG TTGGGAAGCA AGTTCTCTTG
 3551 GCAGGGGGGT GGGGTCTGTT ATAGACCCTT GAGGCCAGG GCGCTGGCAG
 3601 ACCCATCGGG GCATGATGTT AGCCCCGAG TGGAGCCGGC AGCCAGGTC
 3651 TGGACAAGCT GTACCTGTGG CTTCTCCGTC GTCCGACACT CCGTGTGCGA
 3701 GCGTCTGTGA TCCGTCTCTC TCGTTGTCCG TTTGCATCTG GTGCCCCCA
 3751 CCCGCCATCC TGTTACTTTT GCTGTGATGC TGTAATGCCG GGAACGCGTG
 3801 CACACGGTCA CACCAACACT AATAGGACTG TCCTGTCTGC TGTGTGCTCA
 3851 CCACACCCTT TGGGCAAGAG AAGCCCCAC TGGGGTTTTC TAAGGAGAAA
 3901 GGAGGCAAAAT GCTTTTCCGT GTCAATCAGT CCAATCTTGT TTTCACTCTC
 3951 TTGAGCAAAG GATTCTGGAA CCATCTGTCA CCTAAACTTT AACTCTAATC
 4001 TTCTTCTGCT TCCTTTGTCT CTTTCTTCC CTTACCTCGC CCACCCCTCG
 4051 TCTGTGTCCG CCCACCCCTC CTTTCCCTC GTCTCTAACC CCGTGTAAAC
 4101 AGTGGGCAGA CAGACCACCG CTCGGCCAC AATGTCCACC GCGGCTCCG
 4151 GCACCACCAT GGGGCTGGTG GAACAAGGTA GATGTGTCTC GACCAGCGTC
 4201 CCGCCCGCTC CCGCCCGTCC CTCCTGCCAG CATGCAGCCC CCGTGTGAC
 4251 GCAGCGGCTG GCGGGGCTCC AGAGCCGCC CAGAGGCCGC CAGGCCCCCG
 4301 GGAGCCCCTG CTCCTGTGTG GTCACATCCC AGCAGAGCCC ACCACAAGGG
 4351 CAGGGAGGCA GCCCCAAGG CTCCTCGCCT GTAAGAGGAG GGGCTGGGCT
 4401 AGGTGCCCCC TGGGCTACAC CAAGCCCTTC TGGTCTTGGC CCCCAGAGTC
 4451 TGGGGTCCG GAGACCCCA TTAAGAATGG CCTGGGCCCC ACAGGGAGCC
 4501 ACTGGGCTG CTGCTGGGGG GTCTGAATCC TGAAAGGAGA GCCTTGAGGA
 4551 GCAGAGCCAG AGAGGCAGAG GCCCTTGGGG CAGACACACA CCCTGCCCT
 4601 CTGGGGCCGC ATGGAGACGG TGGTCTGTGC TGCTGAGTCC TACACATGCA
 4651 TGTCTGCCCT GAGCATCCCC CCAGGACAAG CCGCTCTGGA GTGGGTGAGG
 4701 GTTTTATGCA CCCTGAGGAG ACTTCAAGG CTTCCTCTTG GGTGTTTCT
 4751 GCAAAGTCTT CCTCCCTGG CCTCAAACCC TGTGAGGGAA AAGGCCGGCA
 4801 CTGGCCACCT GCTCCTCTGG GCTGTGCGGG GCCAGAGCCC AGAGGCCCAA
 4851 GTTGGCTTCT GCCCACCTGC TGGCTTGTGA CCATGGGCAG ACCCATGAG
 4901 GGCTAGGCGA CCCCAGACC TCCTTGACG TCCAGCCTGA GCTGAAGGCT
 4951 GGTGAGAGCT TAGGGCAGGC CAAGCTGACA ACGCCTGGCC ACAGAACACA
 5001 GAGGGCTACA GGGGTGACCC CAGATCCTCC CTGGGCTGAG CTGCTGAGTT
 5051 CCCTGTCTGT GCCTCAACG TGGGCTGGGG ACCCGGCAGA GGTTCAGGG
 5101 TGCTGGAGAC TGCCTTCCCC AGGCCTCCTC ATGACCCACA GGGTGAGCAG
 5151 CCTGGCCTTC CCAGCCAGAG AACCTCCTT CTGGGGAGGC CCAGGGCGTC
 5201 CTCGGGGAGG GCAGTCTATT CTCTCCCAT GAGCCAGTG GACGTGTCTA
 5251 GCAGGCAGCA CCCCAGGAGA GCCCTCCAC GTCTTCTCCA TTTGACAGGC
 5301 CTTTCCAGAG CGCAGGCGGG AGGGGGCTGT GATTAGAAAA GAGTGAGGCT
 5351 AGTGGCTTCT GGGGAGGCAC TGCTGCCAG GGGACAGTGC TGAGAGACAG
 5401 CTGCCTCTAC GCTGCCCTGT GCCCGGGCT CCCGCTGCAA TGCCCGCCTG
 5451 TCTGCAAGTG AACGTGGGGC GACGGTGCAT GAGGCCCTGC ATGTGTGGCT
 5501 CCACCTTGGG CGCCGAGAGC AGCTCTGTCC TGGAGGGTGG TCAGTGATG
 5551 TGGACAGAGC CCAGCATGGC TGTCTTGGGT GACCAGCTAA GGGGACAAGG
 5601 CAGAGGCAGG GCTGAGAGGA CCACCATCC TGCTAGGTCA GCCCAGCTCA
 5651 GCCATATCAC ACGGCAGTGA GCATGGAGCT CAGTTCTCTG CCAATGGCAG
 5701 CTGAGTCTAG TACCATCCAG TCAGAGTCTG GTACCAGCCC ATGTGGCATA
 5751 GCCCCCTCGG CCCGAGAGA GACCCGCTT GTCGAGTGTG CTTAGTTTGG
 5801 GCCTCTGTGG TCTCTCCTGC ATTGATCAGG TGTAAGGGCA TAGGAGACCC
 5851 AGTGTCCGGC CAGCTGCAGG GTGGCAGCAG TTGCCCGGC CTGGAGACCC
 5901 GGGAAATGGG AGTGCCTTCC CAGGATGGAG GGCAGAGGGT CTCTCCTTGT
 5951 CCCACAGAGG CCTGCAGAAC CCCCACCCA GGTGTCTGAG ATGCCTGTGA
 6001 CTGCTCCGCC TACCCTGGGC TCCTGCGGCA CCTAACGCAT GCTTTGAAC
 6051 TGAGACACAG AAAGGAAGTT CCCGTGCCCT TGAATGCTAG TGTAGATGGG
 6101 CATCGACAGG ACTCTGGCCA CCGTGAATCT GGAGTTAGTC CCAGGCAGAG
 6151 ATGTGAAATG AGCAGCCCCC CAAAAATGG TTGGCCGGGA GCCATGCACT
 6201 CAGGAGGGCC GGGCCCATGC ACCCCACACT GCGCCCAAGG CGTGACAAAG
 6251 CGATTGTTTT AAAAGCGGGT TCACAAGGAA GGATGTTTGG GAACTGACTG

FIGURE 3, page 2 of 16

6301	AGACAACAGG	GACGTCTGCT	GCAGGGCTTC	CCAGAGCTCT	GATGGCAGCG
6351	TCGGCCTGAG	TCCCTTCGAGG	AGGGCTGGTT	TGTACGTGGC	ATTTGCTGCC
6401	CACCTGGACTG	TGAACCTCTG	TCTTTTATT	TCCCACTGCT	GCTGTGGTAC
6451	ATCTCCAGTA	GCATAGTTTG	GAAATGCAGG	TTTTGATAGA	CTCAAGGATC
6501	TAAATAGAAC	CCTCTTAGTA	CCAAGGACTG	TCCGGGGTCT	CTGCCAGCCC
6551	CGCCGATGGG	CCTAACTGTG	GTGCCTCCTT	TCTGTGAGA	ATCTTCTGAG
6601	GACATGCCCC	GGGAAAGAGC	TCAGTTCTGC	TGCTGCCTAG	GGTGCCATGC
6651	TGGCCCCGGT	TCCAATGCAG	AGCCTAGCTG	GAAGTACCGC	TGGGTTGGCG
6701	GAGGCTACGT	GCCTGACTGT	CCCCTCGGG	GTGGGGTGGA	ACTAGCCTTC
6751	TGAAACCGCC	TGCTTCAGTT	GGCCACAGCT	TTTTGAAATG	TGTGTTTCTG
6801	GAAGGGACTG	GGTCCCTTCC	TTGCCTGTTT	AGCTCCCCAC	GACAAATGTC
6851	CTCAAGGCGA	GGCTGGATGC	TTCCCTCCTC	AGGCTCCTAG	GAGGAGCCCG
6901	TCCCCCAGCT	GTGTGCGGCA	GCTGCTCACC	AGCAAGGACA	GGATCCCTCA
6951	GCTGCAGCCT	CAGGCTGGCT	GGCACTGGGC	GGGTGTTTCT	GGGATGAGTT
7001	GTGTGTACTG	GAGATGGGAG	GGGAGCTGAG	AGGGTGGGAT	GCACAGACAG
7051	GAGAGGGGAC	TGTGGGGGTC	CTGGAACCCT	GAGTTCCAAG	TCTTCAGGAC
7101	TCCTCCCTCCA	TAGCAAGTTA	CAGGGAAGCA	GATTTGAGCC	ACAGGGAAGC
7151	AGATTTGAGC	TGCAGCGAGG	GGGAGGGTTT	TCAGTCTGTG	CTATAGGGAA
7201	GTGGGCAGTC	GGCATTTCTG	GTCTTGGGAA	CTCACTGGGC	AGGGCTGCCT
7251	TGGGACATCA	GGGAGGTGGC	GCTGTGCTCA	GCTTCACCAG	GAGGGGCCTT
7301	AGGCCTGGGG	ACGGAGAGTG	ATGCCTGAGG	CCCCTCTACT	TCTCCATGGA
7351	TCCTGGGAGG	GACTCCTGGG	CTGGATACAA	AATTGTTGAG	AGTTAAGAGA
7401	TCTGTGAGGA	GAGGGGAGCT	GGGAATAGAA	AGTGTGTGCC	CACTGCACAT
7451	GGGTCCGCA	GGGCCACGTG	CAGCCACTGC	GCAGGCACAA	CCCCAGTCCC
7501	CACAGAGCCC	AGGAGGGGCC	AGAGCCATGG	AGGAGGCAGC	ACTGGGCATT
7551	TGACAGAGGA	GGGGGTGCTC	AGCAGGCAGC	AGGCCCAGGC	CTGTCTATGC
7601	CCTGCGGGGT	GCAGCCTCCT	GATCTCCACG	GCAACCTGGA	GCACCCAGCG
7651	TCAGAACCAC	CGGGAGGGCT	TATGGAACAG	ATGTCCAGCC	CTGCAGAAGT
7701	TCTGGCTCAG	GAGGGCGGGG	TGGGCCTGGG	AATTGTCATT	TCTGACTGTA
7751	CAGGGCGATT	CTGCTGCTGC	TGCTGCTGCT	GGGGTTGGGG	GAGGATCCCA
7801	TTTGAGAAGC	GCTGCAGTCC	TAGGTTGAAA	CGTGCCTGTC	TGTCCCCACC
7851	CAGGCCTGCA	TGGGCAGCAC	GGGATCCCCA	GGCAGGAGGA	CCCAATTTC
7901	TGGCCTGGCC	AGCCAGGGTC	CTGGAGCCAG	GCGGTGGGGG	AGGGATGGGG
7951	GATTGCTGCT	CCACCTTCCCT	TCCCGGCTTG	GCCCGGGGGC	AAGCATCCTC
8001	ACACTTCCCA	TGTCTGTCATC	CCCTTGGCTC	CAGCCTGGCT	GCCTCTCTAA
8051	CCCTGCTGTA	CCGGCTGGCC	GCATGGCCCT	GGCTCTTTTT	GGTGAGCGTG
8101	GTCCAGGACT	GGTGACCTGT	GAGTCCTGGG	CCCGCAGTCT	TGCGCCCCCTG
8151	CCCGAACCAA	CACAAATCTT	GTTTTCTCTC	TCTCTCTTCC	TCCTCACTC
8201	CCTCCCCCTT	TCACCTTTTCC	TTTTCTGTAA	GGTAAGCTGA	CTTCTCTTTT
8251	TGGTTTTTTA	TTTATTTTTA	TTTTTTAGTT	CTGTAATTAA	AATCCTAACA
8301	GCCATGGAGG	GTGTGGGCAC	CGGGGGCTGG	GGCCAGGCCC	CTCTGACCTC
8351	TGAGGGGGAA	TGCTGGGTGA	GGCAGGGGCC	CCGCTGCTGG	GACCAAGTAT
8401	CCTCAGGGGC	TTGTGGGCAG	AAAGGCCTGT	GCTGGCCCCA	GTCAGTGCAC
8451	AGAAAGCGCC	CCAAGGCCAG	GGCTGCTGGG	CAGCTCGGAA	TGAGGGCGAG
8501	CAGGGCTGCC	CTTGGTGCCT	GAGCCAAGGA	GCCAATGGGA	CAGACCTCTG
8551	AGCCTGGGTG	CCAAGTATGA	GGTCTGAGAC	AGGGTGAGCG	CCTGGGCTGG
8601	GACAAGGCC	TCTGAGTGGG	CGGCCAGCTG	CAGCCCCACC	ACCCCTACCC
8651	CAGGAAGGCA	GGGCCCGGGA	GGGCATGACC	TCTGGGGTGC	TGGCTCAGCT
8701	GCCCCACCC	CAACCTGACA	CCGCTAGTCC	TGAGTTCCCA	TCAGGGAGGA
8751	AGCAGCATCC	TGCCTTCCTC	TAGGAAGAGC	TTGCATGTGG	CCCAGAAGCC
8801	AAGGGGGCTC	CCCAGCACCC	ACGGGCATCT	CTGGGTCTGG	TCAGAGGAGA
8851	AATCTGGATG	CTTGAGGAG	CCCCAGGGTC	ATGGAGGAGG	CTGGAGACAG
8901	GGCTGTCTTG	GGGTGATGGG	ATGGCCCCC	CACCTGCTCA	GAGCCAGCCT
8951	GGGTCTGGGA	ACCACCTTG	CCTCAGGACC	CTGGGCTTGC	TCCTGGGGAA
9001	AGAGTGGGGT	CAGGCAAAGG	GGTGGGGTTG	CGCTGCAGCG	AGACCCAGGC
9051	CCATCACTCA	CCATACCTTC	TTCTCCCCA	TGCAGCAGCC	AAGAGTTTAC
9101	TCAACAAGAA	AGCAGATGGA	GTCAAGGTGA	GGCTCCAGCC	GGGCCCTGTG
9151	GTGCCGGGGA	GCCCAGAGCC	TGCAGCTTCA	CCCCACGCC	CTGGGGCTCC
9201	TGCTCTGGAG	TCCCCCTCCC	CCCATGCCCT	GAGAGACACG	GGACAGGGAA
9251	TGCGGAGTGA	GGGGCTTCTC	CCACCTAAGA	GTTCTCTTTC	CCTCTCTCCA
9301	CAGCCCCAGA	CGAATAGCAC	CAAAAACAGT	GCAGCCGCCA	CCAGCCCCAA
9351	AGGGACGCTT	CCTCCTGCCG	CCCTGGTACT	GAGCTCCTCA	AATTCTGCCT
9401	CTCAGCCCCCT	CCTACGCCCC	TGGCTGTGTG	ATTGCCGCTG	GTCAGAGGGG

FIGURE 3, page 3 of 16

9451 GCCGGGTGAA GGTGGGCTCT GCGCCCGCCT GGCTGTCTG ACAGCACTCG
9501 CATGGGCCCC GCGCCCTCATC CCTCACCAGT GGTGAAGTGG AGAGAAGAGG
9551 CCACTGTTGT GGGGGGCTCC AATTGAGACA GGTTTAGGAC TGCTCTGGGG
9601 AGCCCCCTGGC TGAGACCCAC AGATGTTGGG GTGCAGGGGA GAGGCCCAGC
9651 CTCCCACCCA TGTGTACTTG TGGATGTCTC TCCAGGAGTG TTCAGGAAGT
9701 CAGTGAGGCA GAAGATACCC TCTCCCCACC AGGACCCAC CCTCAGCTCC
9751 TCCACCATCC TCAACAGGCC GACCCACAGA CCACTCCGAA GGTCTGGCTT
9801 GGTGGGGCTG GGCAGGATC TGCAGGGGA ACAGCCCATA GTGGCACATT
9851 CCACGGCCCC TGGGGAGACG GGGCCACGGT GGTGCAGTAG AGAGGTGTCT
9901 AAGCCAGTGG CAGCCAAGGG GAGGGCTTGC CGTCACCTCT GTGTTCCCTC
9951 AGTGCTGCTC TGTGGCTGCC TGAGAGGCAG GGCTTAGGGG CTCCCTGCCG
10001 GGGAGGGGAG GGGTCCCCAC CATGCTCCGC TCCAACGCG CCCCTCAGTG
10051 CCCCTTGCCC TGGGGGCTCC TACAGGTGAA CCCTATAGCA GTACTCCCAA
10101 GGATGTAAAG TTGTGGCTGG TGGGTGCCGG CCTTCCTGCT GGGGCGCTGT
10151 GCTGTGTCCC CTCAGCTGTC CTAAGAGCTT TGGGGCTTGC TGGCCCGTAG
10201 GTCCCCATAT TTGCTGGAAG CAGGCTTGGT GTCCCCTGAG AACCCAGGC
10251 CAGGCTTCGG GAGCCAGCCC CAGACCGCCC ACGGGAATAC TGGGTTTGCC
10301 AAATGGCCAC CTTGAGACCC AGGAGAGGAG AGCGGTCTTG GGAGGGGCGA
10351 GCTGCTCAGA GCAGCCAGGC CGTGGCTGGA GGGTGGCTTG GTGCAGCCTA
10401 CCTAGGGCCT TCCAGTGGCC AGGGCAGCCC ACGTGCCAGC CTCACAGCCA
10451 GCCCCATCTC GGACCCTGTC CATCCCATGT GCCACCGCCA CCCCATGAC
10501 ATCTTCAAAC CTGTGCCCCC CACCACGCTG GGGCACAGGT TCAGGCAGTA
10551 AAGGGTAGGG AGAACCCCTC AAGACCGAGC CTGGCTTCTC TGGCTCCAC
10601 ACACATTGTG CAGCTTGTGC GGGCCCCACA CGGTCCATCT CCCACCCTGG
10651 ACAGCAGCAC CTCGCGCAGC CTGGACAGAG CTCCTGTCCA TTCCATCCCT
10701 GCCGGCTGAC CCAGGCTCCT CCCCCAGCTG CTCCACGCGC CCTCCATCCC
10751 TGTCGCCAC TCTGCTCTGC ACTTCTTTCT CGCAGGCTCT GGCCACCCAC
10801 ACCTCCTCTG TCTCCCTGTT CCCCTCCTGG TGGTCTCCGC TTCCTCCTCT
10851 TCTCACTTTC CCTCTCTTTC CTTCCTCTGT GTCTTCTTTC TTCTGTAGGA
10901 GCCTCAAACC ACCGTCATCC ATAACCCAGT GGACGGGATT AAGGTACTGC
10951 CCCACTTTCC TCCTCCCGCT TTCCCCAGGC AGGAGGCTCC AGGCCAGGAG
11001 AGAGGTCTGG GGCAGCATTT GTGCCAGAGT GGAGGGCAGA TGTCCCATGG
11051 CCCTGCGCGC CCCTCCCCGC AGTACGGTAG GGCCCCAGTC CGTCTTCGTG
11101 GGCAACAACA GGACAGACTG GCTCAGGCCC CAGGCGCGCC CCTGGAGGTG
11151 CTTGGCACAG TTGCGCCCGG TCCCCATGTG GCCGACACTC TCAGACCAGG
11201 GCTCTGCGTG TCCACCTAC GGCAGGCACT AGGGCTTCTT GAGGTCTGGA
11251 GCAGGGCCTG CATCTCAGGA GCTGCATCCT TGGCCCTCCT GGCTGTCTCT
11301 CACCCACCT CCCTCAGTG GCCCCAGTG CTTCCTGCTG AGCAGACCCT
11351 CCCTCCTCTG CTCCCTCTCT TGCTCTGGCC ATCAGCTCCC ATCACATTGG
11401 CATCATCACT CTGGGGCCAG GGAAGGGGCT GGCTCTCTGG GGTGGTGGGA
11451 GGGATGGGGC CAGCAGCCAA GCCATTTCCA GGACTTCCAA AACAGCGCCA
11501 CTACACCCA CACGGCCCTC CAGCCAGCT CCCACCTAGG CCTGGGCTCC
11551 TTACAGAGCC CCCAGAGTGC CTCTGTGGGG ACCCCCCACT TCCTTCTGGC
11601 CAGTGCCACC ACCCAGCCCA TCATCAGAAG ACATCTTTCT CCATGGCAGG
11651 GACCAGGGGG TCCAAGGGGC ACCCATGGTG CTAGGCACCA GGGCCTGGGC
11701 ATTCTTCCCA TCTGGCAGCT GGGGATGGGT GCCCTGGGA CCCGTGTGTG
11751 TCTGGGGTGG GTCATGCTCT CTGCAGGACT CCTAAACAAC CTTCTGGGCT
11801 GTGGTGAAC CTGAGCCTGC ACCTAAAAGA CCTGTAGTTC TGGTCTAGGG
11851 CCTCCAAGCA GTGTCCAGGC AGTGTCCAGA CCAGGGGGCG GTCCCCCAGG
11901 GACCTTGTA GATGTTTCTCT CTGAGGAGCA GAGCAGGCCT CCTGGGGACC
11951 TGGGGGATGG TCTTTTGAAG GGCAGCAGCC CTGGAGCAGG GTGGGAGAGT
12001 CTGGGGCCAC CTCTGCCCTC TAAGGCCACC TGAGAGGTGA GGCCGGGGCC
12051 TGACTGGACG TCCAGTCCCA GAGGGGAGG TGCCCTGAGG GAATGTGGGC
12101 GACAGGAATG CTCTGCCTGG GGCCAGGCCA AGGTTCTTGG AGCCCTGTGC
12151 GGATCTGCAG AGCTCCTGGG AACGCCTCAC CCTGTATTTT GGATGACACC
12201 GGCTGCTGCT TCATTGGAAC CAGCCAGTCC CATTTGTGTT TACGTCTTGG
12251 AATTTCAAAA AGCCATTTT CCTCTCTTGT TAAAGAGTCA GCTGAGCATA
12301 CCAGTCTCTC TGCCAGGCTC ATCTTGCTGG GAGAAGTGA GCCCTCATGT
12351 GTTGGGGATG CAGGGTGGCC ACAGCACTAG GGTGGCAGGG CCGGCCTCGG
12401 ACTCCGTGCC AGCCTGTGCT GGCTGCCGTG AGAATGCACC CTGGTGAGGG
12451 GCGCCCTCCC AGGGACCAGC ACAGAAGTGG GTGTCTTCTC CGGTCACTGC
12501 CGCATGAGGT CCACAGAGCT GGGGCCCTGC AGCCGCCAGA GGGCATGTCC
12551 CCTGAGCCCC TGGCCTTTAA GCCCCGTGGA AGCAGCCGAG GCAGAGATCA

FIGURE 3, page 4 of 16

12601	GCTTCAGAGC	CTGGGCTGGT	CCTGACACAG	GCCCAGCCCT	GTCCACCTGC
12651	CCTCAGCCAC	GTCCCACCTA	TCCTTGGCCG	CATCCTGACC	CGCTGCCTCC
12701	CGTGTTCCT	CAGGAGTCTT	CTGACAGTGC	CAATACCACC	ATAGAGGATG
12751	AAGACGCTAA	AGGTACCTGC	ACTTGAGTCC	TTGCCCCCCC	AGCGGCCTTG
12801	GCATTGCTGG	GTTGCTCTTT	GAGGTGGGTG	GGACTTGGGC	AGGGTCAACT
12851	CTCCTGCGAC	GCCTAGTTTA	TGCATGTGTT	GAGGGGCTCA	GGGACCCTGT
12901	AGCTGTAATC	CTGCTCCAAG	CCTGGGTGTC	AGGCCTGCCC	AGAGCGGAGA
12951	AGCATGGCAG	AGATGACCGA	CAGCTGGGCA	GTCTCGGTCA	CCGCATCCAA
13001	GTGAGGAAGC	CACGGCTTTG	CATGGAGGCA	GGTTCTCCAC	ACCAGGACCC
13051	TCACGGGGAA	ACAGGCCCAT	GGGTAGAATT	TGTTCCAAGA	TGCTGTCTTT
13101	GTCTTAAAGC	TCCTTAAAGC	TGCGTTTCTG	TCCAGCATGC	ACTTGCCAAG
13151	TGGCCGGGCA	GCTGGGTGAG	TGTTTCCGTG	TTTGCCTTTG	CTTAGCCAGG
13201	AGTGTCTCTG	TGCGGTGGGT	TTCTGCACCA	CAGATTCCAG	GGCCCCCTCC
13251	CTTGCTCACC	CAGGCCAATG	TCTTGTGTGT	TCCCCAAGAG	GCCCCCAGGG
13301	CACCAGGCAC	TGGGTCATGC	TCCATGGATT	CTGCCGCCTC	CAGACCACCC
13351	ACATGGGGCC	TGCTGACCCT	CATCGCTCAC	ACGGTCACCT	AATAAGCCTT
13401	ATGCTGTTCT	CAGGGCTACC	CTGGTGCCCA	AAAAGGGTCA	GCCACTCTGC
13451	CAGTTTAGGG	GAGAAAACCT	CTCACCTGTC	CAAAGCATAG	CCTTGCTCCT
13501	GCCCGGCCCTA	CCCAGCTATG	ACACTGTCCC	TGAGCAGAGA	TGAGCACAGG
13551	ACTTTGGGGC	CTGGATGCCG	GAGAGTGGGT	GTTTGTGTGA	TTCCCCCTGCA
13601	GTCTGGAACA	GGCCCCAAAG	GCAACAGCAT	GAAGGCTGTC	CAGAGGTTCT
13651	GCATCACCTT	CAGCCGAGTG	GGGTGCTGAG	CAGTGAGGGA	GGGGACCTGG
13701	GAGGGGGGCC	CAGCCTGAT	CCTGCAGGGG	AGAAGAGAAG	ACAGCCAGAA
13751	GCCAGCAGCT	GTGGCTCAGA	TCTGAGCCCG	AGCAGCCTCT	CGAGGTGGAG
13801	GCAGACACCC	CCCACCCAC	CCCGTGACAG	AAGAAGCCTT	GCCAGCCTGC
13851	CCTGAGGCTG	GTACAGAGTC	CAGGCAGGCT	CAGTGGCCAT	CATGCCCTTA
13901	CGATGACTGT	CACTCCCTCT	CCGTGCGCCT	GGCCTCTGCT	GGCTCTGGCC
13951	AGGGGTGGTC	ACAGCACTAG	GGTGGCAGGG	TGGCCTCTGA	CTCTGCGCCA
14001	GCCTGCACTG	GCCTGTGCTG	CCCTGGCCTC	TGCTGGCTCT	GGCTCTGGCA
14051	CCGGTCCCCT	GTTGGCTCCT	TCAGCCTTCA	CATACCTGCT	GCGGCCACCA
14101	CAGGCCCAGG	ACCCCCACAG	GGTGGCCACC	CCACCTCCAC	CCCAGGAGCC
14151	CCAGGTATCC	AGCTGTCAAC	CCCTCCCTCC	CTCCTGGCCT	CCCCCTGTCC
14201	TTCTCCAGTT	GCCTTCTTTT	CCTGCGGGCG	CACCACCCAC	CTGCCTGCCT
14251	CACCTGTTC	GCCTCAGCCC	CCAGGTTCCC	CGACATCTTG	AGCTCAGTGA
14301	GGAGGGGGCTC	GGGAGCCCCA	GAAGCCGAGG	GGCCCCTGCC	CTGCCCATCT
14351	CCGGCTCCCT	TTAGCCCCCT	GCCAGCCCCA	TGTAAGTAGC	CTGGGTCTTG
14401	CTGCTGTGGG	GGTCATGTTG	GAGGGCTGGC	AACCCCTAG	AGGGGCCACT
14451	CCAGAGCCGA	GGGCAGGCTG	AGCGTGACCC	CTGGCTCCAG	CCTCATCACC
14501	CCACAATCCC	TCACTGGGGC	TTTCCAGGGT	GGCCCCAGCC	CATCGAGCCC
14551	CACCTCTTTG	TGAGGAGGGC	CCTGGACCAC	TTTCTGTCTC	AAGGCCACTG
14601	GGCAGGATGG	GAGGCCCTGG	AGGCTCGGGC	CTCAATTCCA	GTCTTCAGGG
14651	TCGGTGACAG	CCTCACTCCA	CCTCAGCTTG	CGGGCGGGGG	GGCTCCCTGC
14701	TATTGAGGCA	GGCTCTGATT	CAGGGCCTGA	TCCCAGGGCC	CAAGGGGTCT
14751	AGAACACGGG	ACCCCTCCCA	CTGGCCTCCT	CCGCCTTGCC	GCCGCCTCGT
14801	GTGTCTGTCT	GCCTCATGTT	CACGTCTCAT	CTGTTCCACC	CCAGCCCCCA
14851	GGGATCTCTG	ACATCCTGAA	CTCTGTGAGA	AGGGGTTCAG	GAACCCCA
14901	AGCCGAGGGC	CCCCTCTCAG	CGGGGCCCCC	GCCCTGCCTG	TCTCCGGCTC
14951	TCCTAGGCCC	CCTGTCTCTC	CCGTGTAAGT	AGTGGCCCCC	AGGCCTGCCG
15001	CCTCTGCTGC	CGGACAGCTC	CCTGCGAATG	GCCGGCGCTC	AGCAGCTTCC
15051	CACCTGCATG	CACGGCCACG	CTACCCTGCC	CCGGCGCCGC	AGCCTGGAGT
15101	CCTGCCCTGG	CGGGGCTTCC	TGTGGGCTCC	CATGCTAACC	AGCAGGGCAG
15151	CTCCTGGCTT	CTCCCTAAGG	GGCCAGACC	CCTCCACGGC	TCCTGCTCCC
15201	ACTGCCACTC	CCCGCTCGCT	GTCCAGCCCC	AGGCCCTCT	CCAAAATGTC
15251	TGTCCACGCC	CTGGGCAGCC	CTGGCCCTC	CGAGGCCCCC	CATGCCCTTA
15301	GGCCCTCTCT	GCTGATCACT	GTCCAGCCCC	CACAGACTTC	ACACCCACCC
15351	AGGGGCCCTG	CCCATGGTGC	CCAGGAGCTG	CACTCAGGGC	CACCCTGGTT
15401	CCTGATGTGG	CCCCAACCCC	TGAGCACCCCT	CCCTCAGTCT	AGGAGGCTGA
15451	GGAAGGTGCC	AAAACCTGGAA	CCCCGACCAG	GGTCTCTGGA	GCTCACCAC
15501	AAGGGGATAG	TACGGAGAAT	CATAAGCCTG	GCCTCTGCTG	ACCTGGGCTG
15551	TCCTCATGGG	GCCAGGCCAG	GCCTCCTCTG	TAACGCCCGT	GACTCCCTCC
15601	TCTCCCTGTA	ACCCCGTCCA	GCGTTCCTCA	AGGGCCACTT	ACCTGACAGC
15651	TTCTTGCTGG	CCAGCAGCCT	CTCCCTGGAG	GGTGCCCTCT	GCCCCAGCA
15701	GCTTCAGCCC	ACGCCACCCG	ACAGCCAGAG	CATCTGCCCT	TCACTCCTGC

FIGURE 3, page 5 of 16

15751 AGCCTCCTCT CCACGCACCA CGCTGTCCGC AGCAGCACCC TCTGTCCCCC
15801 TGTCCTCCCTC CGTCCCCCA TATCCCCCTC GGTCAGCCTA CAACCTCTCC
15851 AAGTCCCCCTT AAGTCCACGC TCTATCCCTA CATCCCCCTC TGTCCCCCAA
15901 ATTCCCCCTCT TTCCCTCAT TCCATTTTCC TCCCCAACT CTGCTCTGCC
15951 CCTCACATTC TCCCTCTGTC CCCCACACCC TCCTCTGTCC CCCACACCCT
16001 CCTGTGTCCC CCACACCCTC CTCTGTCCCC CATATACCCC TCTGTCCCCC
16051 ACACCCACCT TGGTCCCTTG CACGCCCTTT TCTGTCCCCC ACACCCCTC
16101 TGTTCCTTAC ACTCTCCCTC TGTCCTCCAG ACCCTCCTCT GTCCCCCACA
16151 CTCCCTCTGT CCCCCACACC CCCTGTCCCC CACACTCTCC CTCTGCCCCC
16201 CAGACCCTCC TCTGTCCCTT ACACCTCCCT TGTCCTCCAT ATCCCCCTCT
16251 GTCCCCCACA CCCTCCTCTG TCCTCCACCC CCTGCCCCC ATACCCCTT
16301 CTGTCCCCCA CACTTCCCTC GTCTTCCACA CCCCCTCCTG TCCCCCACAC
16351 CCCCTCTGTC CCCCAGACTC TCCCTCTGTC CCCCACACTC CGTCTGTCCC
16401 CCACACCTCC TGTCTTCCAC ACCCCCTTCT GTCCCCCACA CCCCCTCTGT
16451 CCCCATACT CTCCTCTGTC CCCCACCTCC CCTCTGTTC CCACACCGCT
16501 TCTGTCCCCC ACACCCCTC TGTCTTCCAC TTCCCTCTG TCCCCACAT
16551 CCCCCTCTGT CCCCCTGACC CTCCTCTGTC CCCTGCACCC TCCTCTGTCC
16601 CATGCACCTC TCTCTGTCCC CCACATCCCC CTCTGTCTC CACACTCCCT
16651 CTGTCCCCCA CATCCACCTT GGTCCCTCA CGCACCCCCA TCCCCATGA
16701 CCCCTTCTGT CCCCCACACC CCCTCTGTCT TCACACCCCC CCTCTGTCCC
16751 CCACACCCAC CTGCTCCCC TCATGCCCC CATCCCTTAC ACCCCACTT
16801 TGTCCCCCA CATGCCCTC TGTCCCCAC GTTCCCTTCT GTCTCCACG
16851 TCTCTCCAT TCCCGTTTC CCTCTCTGTC CCCCAGCTC CCCTCCATCC
16901 CCCACATCCC CTTCTTTTCCC CTATATCCCC TCTGTGCGCC CAGGTCCACC
16951 ATCTTCCCCC CACACCCCCC CATTCTCCCT TCCTCCCCTC TGTCCCCTTG
17001 TGCCCCATCC CCCACATCTG CCTCTGTGCC CCTCAATCTC TGGCTTGGCT
17051 GTCTGCCCAT GGTTTCTCTC CTGCTGTCCC CCGTGCCTG CCTTGTGTTC
17101 ACGTCTCGTC TGTTCGCCC CAGCCCCAG GATCTCTGAC ATCCTGAACT
17151 CTGTGAGGAG GGTCTCAGGG ACCCAGAAG CCGAGGGGCC CTCGCCAGTG
17201 GGGCCCCCGC CTGCCCATC TCCGACTATC CCTGGCCCCC TGCCCCCCCC
17251 ATGTAAGTAG CACCTTGAGT GGCCGTGGCA GCGGCTGCCT GGAGGGGCTC
17301 GGGGCGTGCG AGCCTGGCAG TGGTGCTCTG GGAAGGGCCA TTCTTGCGGA
17351 GGAGGGCGGG GCACAGGATC CCTCTGCTGG GTCCCAGGGA ATTGCTTTGA
17401 AGCACATGAA GGTGCCACTG GGTCTCAGAA AATGGAGGTT ATGGTTATGA
17451 AGTGTGTATG ACATATGTGT ATAGGAAGAG CGTCCGAAAG AGCAGGTTTG
17501 TTGCCGAGCC CAGCATTCGC AACCCTGAGG TCCACAGCTT TCTCCTGATG
17551 GGAGGGGAAT GGGTGGCAA GGGTCTGCGC GTGTGGCAAG GGCTAGCACG
17601 CCAGGAGCTG CTGGCTTGGG TCAAGGTGGA CCTGCTGGGC CGGGACAGAA
17651 AAGTGTCACT CCCGCTCTGA GACGCTCTAG CATTAGAGCT GTCCAAGTCC
17701 AGACAGCAGG GAGCAGGTGG GGATCGGGAG GCGCGGATCT GGGGGGCAGC
17751 TGGGGCCAGG CTGAAACAGA GCGGGCGGGA CAGGAAGCAC AGGCTGGGCA
17801 CCCTCCCCCG CAGGGGAGGA GCCAGGCTGG GCCACCTCCC GGTCTGTCTG
17851 CCGACTACCC GCAGTATCAC TTACAGGGAT GGATGACATC CCAGGGCTGC
17901 TGCCACCCCC ACCTGTGGGG AGACACCAGA CTGGGGGTGG TGTGGAGATA
17951 CTCTTAGAGA AGAGGCTGCT GGGCCACGGG CTCGCGATGG CAGGGCAGTG
18001 GCTAGGTAAG TACTTGAGGG ACAGGTGGGG TCTGCTTGCC ACCGTCCCCCT
18051 CTGCAGGCTG GGCCTGGGG CTGCTGCAGG CGGCCAGGGC AGAAGGGTGT
18101 GGGGAGAGTG AACCACAGG AGCAGCGGCT CGAGGAGGGG GATGCAGGCT
18151 GCAGGCTCAA AGGGGCACTG GATCCACCCT GGGTGCCCGA GAGAGCAGGG
18201 GGCAGCCCCC GGAGGGGTAC TCACCCCAG AGCTTCTGTG GTCGGCTGAG
18251 GACCCCAAGC AGGGGTTGAC TGAGGGGATC AGAGGCAAGC AGCTGAGGGG
18301 AGAGGCCAGG TTCTTGATGC TGATAGGGTC GGGGTGCCTG GCGCACCAGA
18351 ACTCAAGGAG GGAGGCATGG GGAGGGGCCG CCGTGCAGCT GGGGTGGGTG
18401 CACCGCAGAG GCTCTGGAG TGGTCAGAAC CCCCAGACCC TGCCACTTCT
18451 ACAGCAGCTC ATCTGATTTT AAGGGGCTTG CTGCCCTTGC AGAAGTGGAG
18501 GGGTGTGCCC AAAGGAGCCT GCCTGGAAGG TCACCCCATC AGGTTGGCAT
18551 GACCCAGGCC CAGGACTGCA GCCTGCCCTC AAGGTCTGTG CAGTATCTGG
18601 GGTGAGTCCT CTGAGGACAG GGCCAGGGT GGGTGTGGAG TGGCCAGCTC
18651 GGGGCTCGGT GTCCAGGCTC ACCTTCAGGG GCCACAGCAC AGACCTGCCC
18701 TTCCAGAGTC TTCCCTGAGC TTGGCTGGGG AGGAGGGGGC TGCAGGAAGG
18751 AGCTGTGAGC AGGGCAGGAT GGAGATTCTG GTGGCCCTCC TGGGAGGGGC
18801 TGGGCAGGGC TGGGAAAGGG GTGGGTGAGA TGTTCGGAA CTCAGGGAAA
18851 GGAAGATCTT GGGTACTGCC CTGGGGGCAC CTGGGCCAG GTGGCAGGTG

FIGURE 3, page 6 of 16

```

18901 GCCAGCTTTC TGCCTCCTTT CCACCTCCTT TCTCCAGAAG GCACCCACCA
18951 GCTGTGTAAA TAGGGCAGGT GCCCAGGCC CGCCTCAGGC CCCGTCTCCT
19001 CCCACCCAC GCTCTCTAAT CGCGGATTAT ACACAATCCA GCCTGATCCC
19051 TGGGCAGCTG CCCTCCCTCC CGCAGCCACC TCTGGCTCTG AGAGATGGGC
19101 TTGGGGCCAG CCTGGGGTCC CAGGAGTCCA GGCCAGGATG AGAACCTGCT
19151 CTGACCCAC CTGGACGCAT TAGGCCTGCC TGGACCTGTT GCCTCACCCC
19201 AAGAGAGCCA CAGGCAATGC AAAGGCTCCT GTTCATGTCA GGGCACCTGG
19251 AAGGCCTGAC TTGCAGAGGC TCTTGGCTCG TGCAGACCCC TCCAAGCCCA
19301 GGCCCTGCCC ACCACCTCCC CTTTGTCTCT GGAAGTCCCA GGACAGCTTG
19351 TCCTCAGCCA GCAGGTTTCC CGACCCGGGC ACCTCTTCAT GTTGGGCCCC
19401 CCTCCTTTCC CTCCATCAGG GATCATGCCC TTCTTCAGGG GCCTGGATAT
19451 CAAGGACACA AAAGCTCCCA TGTGCTATGT GGGGAGGCAG AGTGGGGGCT
19501 GGGTTGAGCT GGGGTCTGGG CAGCGCCATT CCGCAGGGCA GGGGCAGCCT
19551 AGGCTTCCCA TCTGTGGAAT GGGTGGGTGG GTCTCACAAAC GGACCTGCTT
19601 CCCGTACTTC AGCACGGTTA CCACTCTTGA TTGGAAGTCT GACCATGCAT
19651 CTCCTCTTCT GTTTACTTCA CGCTTCTCT TCCCATCAAC TCCCATTTTA
19701 ATTACAATTT GTTTAAAAGC ACTGCATATT ACTTCATTAA ACAGAAGATT
19751 AGTTTCACTT ACCATTAGTG TAAGGTGACT ATAGAACCAA AGCAGACTGG
19801 AAACCAAATG ACATAATGTC ATTCTCTTCT CCATTCCAGC TGCCTGCTGC
19851 TGTGCGCCTG AGAACCCCTG TGGAGTGGGA GGGGCAGCTG TCTCTGTACA
19901 TTAGAAAGGG AGGTTAACTA AGTGACAGGA GGTGTTTGGG ACATGTGGAC
19951 ACCAGACTTC TCTCTTGATG CAAGGAGGGC AGAGCCAGGC AGCCTAGTGG
20001 GGGCTGGCTT GGGGGCTGCT GGAAGGACTG GCTACAGGTG GAAGAGAGGT
20051 CAGACCTGAA GCTTGGGGCC ACCTCCAGGA AAGGACAGGT GAAAGTGGAG
20101 GCATGAGGCA GGGGAGAGGC AGGTGCCAGG CAGAGGGTGG AGAGGAGGCA
20151 GGAACATAGC AGCTGGGGCG GGGGCGGGCC CTCAAGTGTC ATATGCTACT
20201 TTCTTGGGGC CCAGGGGCAA GGACAGGAAC AGCCACAGCA TGTGTTGGGA
20251 CAGAGCCCTG TGCCTTCCTA GAGCTGGGCA GGTGGAATGG GGCAGGAATG
20301 GGACTCGTGA TGGCTGCAGC AGGAACTGGA GGGGAAGGGG CTTCTGGATC
20351 CTGCAGCCTA CCTTCCTAGA GGCCAGCTTT CCGGGGTCCA CCAGGTGGGT
20401 GGGAAGTGGG CTTGTGTAGC AAGACTGCCC TGAGGACCAT CCATGACATG
20451 GTCTAGATGA AAGTTAGGAA AGAAAGGGAG ACAAGCTGGC AGCAGAAGTA
20501 CAGCTGGGTC AGGAGCAAGG GCCTTTCCAG ATAGGGACAA CCCAAGAGTG
20551 CACATGTGCC CACGCCACAC AACACAGGCA CACACGACAC GTGCACGCTC
20601 ATAGGCACTG CACACACACA TGCACAGGTG CTATGCATA TGTATGAGCT
20651 TCATCTACAC ACATTACAT GCCGTCTGTC TTATGTGCAT GTTTCCATAC
20701 ATGCACATGA ATGCACAATC ACGTGTACAC ACATGCATGT GATCACATAC
20751 ATGAACATGT GTGCACCCCA CTCCTCAGGT GCCATCGGGC TCCTCCTGCT
20801 GTCACTGTGC AGCAGGGGAC ATGAGGCCCC AGAGCAGACA GGTGCAGCAC
20851 AGGCGTTCCC AGGCAGTGCC CCACACACAT GCATGAGCAC ACCCGGGCAT
20901 GTGGCGCCTC CTTTGTGGAC TCAGTCCACC TGCCAGGTGG GCTCCCTGGT
20951 GGTGTGAGCT CCCAGAGGTC TGGCGAGAGA GATAAAGGCA ACCCCACCAC
21001 CAGGCGTGCT GAGAAATCCC TCTTCTGGCT GGGCACAGTG GCTCATACCT
21051 GTAATCCCAG CACTTTGGGA GGCCGAGGTG GGCAGATCAC TTGAGGTTAG
21101 GAGTTTGAGA CCAGCCTGGC CAATATGGTG AAACCTCATC TCCACTAAAA
21151 ATATACACAC AAAAAAATTA GCTGGGTGTG GTGGTGTGCA CCTGTAGTTC
21201 CAGCTACTCG GGAGGCTGAG GCAGGAGAAT CGCTTGAACC TGGGAGTCAG
21251 AGACTGCAGT GAGCCGAGAT CATGTCACTG CACTCCAGCC CGGGTGACAG
21301 AGTGAGACTC CATCTAAAAA AAAAAAGAA TTCCCTCCTC TGGGAATTTA
21351 GACCACAGAC AGGTTGCATG TATGTGGCCG TTGGAGGCAG CACTCACAGC
21401 AAAGAGTGGA AACGTACCA CAGGGCCTGC CTCTCGGTGA AAATGGTGTC
21451 CTGCAGGGCG GGCAGCTGTT TGAGGGCAGG TGTCACAGGT GCGGCCTGCA
21501 GCAGCCTGAG GGTCAACAGG CGCAGTGCTG GGAGTGACAG GACTTCCCCC
21551 ACAGGGAGAG TCCCAGGAA CCTGCTTCCG GTGCACTTCT GGGGGTTTGA
21601 GTTTTTTCCA CGGACGAATT ACTTTGAGAA ACCACTGTTA CTCGTGTGTA
21651 TAGGTGAGCG TGCGTGTGCA TGTGTGTTCT GTGTGTGAGT GTGCATGTAT
21701 GTGCGTGCCT GCGTATATAT CCTCGCAGAT ACGGCTAGGG ACCTCACTCA
21751 GGACAGTAGT TCTGCCTGAG GAGAGTGAAT GCGGCAAGAT TGAGGAGAAC
21801 ACAGGCATCT TCAAACATA TGTGCGGTGC TTTATTTCTT TAAAAATGCG
21851 TCTAAAGCAA ATAGGAAAAT GTTAAGATTT GAATCCGTAG AGTGTGGGTT
21901 CTATTATTCT TCCACATCT TCCATACGTT TAAATCTTT TGCAATGAAA
21951 ATAAGCTGTA GTTAAAGCAG CAATGCAGGC TGCCAGTGAG CGCCCCGGAG
22001 GCCAGTGAGG ACCAGCATGG CTGGGTGGCC TGTGGAATC CAAGGGGGGC

```

FIGURE 3, page 7 of 16

```

22051 GGGCAGGAGC TGCAGGCAGG CGCCCGGGAG TAGCCCGGGC ATGGGGGTGC
22101 GGGGCAACAG GGATGTCTGC AGGGGTAGCA TGTGGGCCCC GGAAGCAAG
22151 CAGGTGGAGC CAGCCGGATG CGGCTCCTAT GAGAAAAGCG GGAACAAGA
22201 GACCACGCTC GTTCTTCCCTG CTGCGGGGAC AGCCCTGGTC ATCGCTCCGG
22251 GGAACCTGCT AGCCTGCGCC GCACGTGGCC GCCCCCTGCT GCTTCTCCT
22301 CCCC GGCCCTC CGGGTGGCCT TGCTGACGGC TCCTTCTCTG AGGCAGGTCT
22351 CTGCTTCTC GCCTGGTGCC TGCACTCAGT AGCCCCCTCA CCAGAGCTGC
22401 TGGGTGAAGG AAGCACTAAG AACCAAGGC TCGGGAGGAG AGTGGGGCCG
22451 GGAAGCTGCA GGAAGCGCA GGGCCAGGCC TGGTGGGCCC AGGGGCTGGC
22501 TCACGGGAGG GCAGGAGGGA GACTGTGGCG GACAGCACGT GGGGCCAGGA
22551 GGTGACCTCC AAGTGGATTG TGGGTGGGTT TTTTGTCTC TTTCTGCATT
22601 TTCCAGGCAT TTTGTAATGT GGATAGAATA TTTCTGTTCT TCAAAAATAC
22651 TTTAGTTAAG AAAAATAAGA TGAAGCTGT TGCACTTGAA AATGAGGAAG
22701 CCACTGGTGA TGCAGGGGGG GCGGCGGAGA GGACCTCTTC TGCAAATAGC
22751 GGCAGGAACA CGGCATGGAT GCAGCTCGCG CTCCCCCAGG CCCTCCCCTG
22801 GGCTGTGTGG AGGGGTCCGG GGGGAATGGG CCAGCGCCCA GTGGTCACCT
22851 GGCATGTCT CCCCACAGCC CGGAAGCAGG AGATCATTA GACCACGGAG
22901 CAGCTCATCG AGGCCGTCAA CAACGGTGAC TTTGAGGCCT ACGCGTGAGT
22951 CCCTGGGGCT GGGGGGGGGC TGTGCAGGAC AAGGATGTGG GACCCTTGGG
23001 GGGCCTTGCT CAGAGTCAGG GGTCCACGGG GCCCTCCTC ACTTGATTT
23051 GGCCCCCAGG AAAATCTGTG ACCCAGGGCT GACCTCGTTT GAGCCTGAAG
23101 CACTGGGCAA CCTGGTTGAA GGGATGGACT TCCACAGATT CTACTTCGAG
23151 AACCTGAGT GAGGAAGCCC GGTGGGCAT GAGGGGGCGG TGCCCCCAGG
23201 AGAGCCTCTC GGCCCCTCCC AGGGACAGCA TGGTGGCTGC CTATGGAAGC
23251 CCTGTCCCCT CTGTGCCCAG GGTGGCCAG CCACCTCTCC CCCGCCAGAG
23301 GCCATACCCA GCCCCCAGAA TCCCCTCTT GGAGGGGGCC ATGCTGCTCC
23351 CAGGAGAGCC GAGCCTCCCC AATAAGGGGA GTTGAGAGAG GGAAAGGATT
23401 AGGCTGGTGG GGTGGAAGAC GGGCACCAGG GCAGTCATGG TAACCCGAGA
23451 CCCCCGCTTC CCGCTGTGTC CACAGTGCTG GCCAAGAACA GCAAGCCGAT
23501 CCACACGACC ATCCTGAACC CACACGTGCA CGTCATTGGA GAGGATGCCG
23551 CCTGCATCGC TTACATCCGG CTCACGCAGT ACATTGACGG GCAGGGCCGG
23601 CCCGACCA GCCAGTCTGA GGAGACCCGC GTGTGGCACC GCCGCGACGG
23651 CAAGTGGCAG AACGTGCACT TCCCTGCTC GGGCGCGCCT GTGGCCCCGC
23701 TGCACTGAAG GTGAGTGTTC TGTGCTAAGT GACAGCTGGG GCAGAGGGGT
23751 GGGGTGGTGT TGGGTGGCTG CAGCCTGGGG AGGCGATGGG GAGCGTGGG
23801 GCCTGTGGCA GAGCCCATGC CTGGGAAGTC CCTGAGCTTT CTGTGTGAGG
23851 CCACAGGAAT GATGTCAAAT TAGGGACCAC GGCAGGCTGG GTGTGGCAGG
23901 CCTCCCCAGA GGAAGTGGGA GCTGTGTGAG GCCTGAGCAG TCCACTGAG
23951 CCAGAGCTGG GTGGGTGCA GGTGGATGGG CCCCAGGAG CACAGTCTG
24001 GGCACCATGC CCTGTTTGTG AGGACTGTTA GAGCCCCAGA TGGGCGTTCC
24051 CCAGGTGGTG GGTGCAGCGG GCCCAGAGCC CAGTTTTACA GGGATAGTAG
24101 TAATTGGGTT GGGCACCTTG AACCTCTCTC CCGAGTGGGC CTTTTTCTGG
24151 ACTTTAACCC TCTCTGCAGT GCCGCATGGC AGACAGCAGA GCCTGGGGGT
24201 GGATGGGAGA GGGGGCTGCT GAGGAGCTGA CCCACCCGCC CCATTTTACA
24251 GCTGCGCCCT GGTTCGCGG GACAGAGTTG GTGTTTGGAG CCCGACTGCC
24301 CTCGGGCACA CGGCCTGCCT GTCGCATGTT TGTGTCTGCC TCGTTCCCTC
24351 CCCTGGTGCC TGTGTCTGCA GAAAAACAAG ACCAGATGTG ATTTGTTAAA
24401 AAAAAAAAAA AAAAAAAAAA AAAAAACAAG ATGACGACGA CAACCACAAA
24451 AAAAATTGAC ATCAGATGAA ATGAAAAAAA AAAAAACAA AAAAAACTAA
24501 AGGAAGGAAA AAGCTGTAAA AATCACTGGC ATTCGTGGGG CCACTCCCCA
24551 CCCAAGCTCC ACGTGTGTCC GTCTGTGCTC CTGGCCTCTG GGGGACCAGC
24601 TGGGACATGA ACTTGTCTGC CAGGCCCCCG TCGCGTGTG AACGGTGTTA
24651 GTTTGTAGGT AACGCACACA CCCACACCT AAGGTGTCTG CATCCTCTG
24701 CCAACGCATG GGCTCCACGT GGTGTGCTCG CTGGCTGTG TGACTGTCAG
24751 CTGTCTCTTG GGAGG3GCTG TGGGGGCCCC CTGGGCTGCC TCCTTTCCCG
24801 CTAGTTGTGC CTGAGAGTTG CTGTTGTTCC TGCTTTCCCT TCCCTTCCTT
24851 TCATCCCCCTG AAGGGCTAGG TGTGGGTTTT CCGTGCCCGG TATCCCCACA
24901 CACCAGCAC GGACAACCTT TCGGCAGAGC CCAGGCCGGC CCCTCACCCC
24951 CTGGAGTATT GAAACTGGAG TCCCGTCCCC AAGGCCTTCA GAGATGCCCC
25001 TACACACCCA GGGCTCCAGC TCTGTTCTTT CTGGGGGAGT AAAGTGCAAA
25051 GAGGGGCACA GCTTAGTTTT GGGCCTCTCG CCGAGCAAGA GACAGCACTG
25101 CTGGCTACAG CTCCAACACA GCCAGCTGTG GCAAGAGGAC TCTGCCTGGG
25151 CTGGCCCCCC TCCTGTGTGA GGTGTCTGTC CCTTCTCTGC TGGCCAGCAG

```

FIGURE 3, page 8 of 16

25201 CAGATGCACT GGCAGCTCCC AACCTGTTT CCGCCCCTCG GCCCTCCCCC
25251 AGCCTGTTTC GCTTCTCTGC AGCCCCAAG GGGGAGCAGA CTTTGTGACAA
25301 AGGACTGCGG GCCTCGCTCA AGTCCCTGAG CCCCCAGCTG AAGCTGGGAG
25351 GGGAGGCCAG GCTTTGTGTC TGGGCATATT CGTCTGCTGA TGGGGTTTGG
25401 GGAAGCCTGG GGCTTGGGGT TTGGTCGGGT GGTGCAGCTA GTGGCAGAGC
25451 GGGATCAGAG GTGGTGGCTG CCCAGCTTCT GGGCTGAGAC AAGGGTCTGT
25501 GCAGGGGTTT ACTGAAGTGG GAGTGCCTTT GGAATCTGGG CCGGGAGCAG
25551 AAGGGAGCAA AAGCTACAGT GGGAGCCAGC CTAGGGCACA TGGGAGGCGT
25601 GAGGGCAGTG CTGCCCCGTC AGTGTGAGGT GTGCCAGTGC CTTGGCGGGC
25651 TGCAGTGCGT GTGAGGGCAC CTTCTAGGTG GGCCAGGGAT GCAGCTATGG
25701 AGATAAGGCG GGCTGGGGAC AGAAACAGGT GGGCACAGGG CCCAGGACAC
25751 CAGCGGATGG AGGGCAGGGT CTAGCCCTGT GCTCCTGAGC GTCGGCTCCC
25801 TGGGTTTCGAG GCGGTGGGTC CCCGGCCCCC TGTGATGGTG TGTACCATGG
25851 GGGAGCTCGG GGACAGGGCA AGCCCAGCA TGGTGGGGCT GCAGGGTGGG
25901 TCTGAAGCCA GGTGGGGTGG GGGTGGTCAC AAGCCCTGAC TGCAGAGGGT
25951 CAGGGGCTCC CTGCCCCAGT CCTGCCACT TTCAATTAC ATTGTTTCA
26001 ACAAGGATTT TCTTTATCTT CCCCTACAAA TCAAGCCAAG GGAGGGGCAC
26051 AGAATGGGGA ACAGGACACA GGATCCTAAA CTCCAAGGGG ACTGTCCACC
26101 GATGAACACT CAGAGTGGAC ACCATCTTCC GTCCACGCTG TGCCAGGAC
26151 AGCTGTCCCC ATCCATGAAC ACAGGGTAAA CATCTGCCGG GCTCCGCACC
26201 AGTGGCTCCC TGGGCCATGG GACAGCGGCA GGGCTCACCA CGGACAGCAC
26251 GTGGCCCAGC AGCCGGCCAC CCTGGCGTCC TGGGGCCTCC TCCCCTCCTC
26301 TCCCTCTCAC CTTGTCACTT CCACGAGCT GCCTGTCTGG GATAATTGG
26351 GGATTTTTTT TCTGGGGGAT AATTCTTTTG CATGACCCCT AAAGAGCAAG
26401 CCACACCGGT CTGCTAGCTA GGTGTCCGCG GTGTGGTGGT GCGGGCCGCT
26451 GGCCAGCGCT GCAAGGGGTC GGCTGCCAC GGTGCTGGCT GGCCTCCCTT
26501 CCTCTCTCTT TTTGCTGAGT TTCATTGTCT TTTCTTTCTG AGCCTTGTA
26551 GTGTACAAAA ATTATTCTTA TTTTGTCTG TCTCGGAAA CTGCAATAA
26601 AAGAAAAACA GGACAAACTG CTTCAAGTGC AGCTGGGTGC TTTAGCTGGA
26651 ATCCTGCCGA CCTCCTGCGC CAAAATACAG ACTCAAGCCC GGTCCCTGGC
26701 CAAGACCCTA CTTGGGCCCC TCCTCCAATG AAAGGTAGTG CTATGGGAGC
26751 CCTGAGCTGG CCTGACAGT CCTGAGCCCC TCTAGGGTGA ACGGCTCACC
26801 CCAGGTAGGG CACTAGTCAT AGATCATAGC TCTACCAGCT GTCTCCACCT
26851 CTTCTCTGTT TCCTCTGAAG TCTTCTGGG CCAGCGCTGT CCACCCTGAA
26901 TGCTGGAACCT GAAACTGGAT CCCAGCCCCC AACACCCCTG ACCTCTCCAT
26951 TCACCCCCGG TGGCCGCTAA GGATGTGGCC AGGGCAGCCT CTGGGCAGGA
27001 AGGAGCCCCA GGACCAAGAC CTCTGGCTGT CCTGCTGTTT CTTTCCGCCC
27051 CTGCTACATG TATTGGCTAT TCTGGATGCT GAGGACACAC AGTGACCACA
27101 GAGCCGGGCT CCACCCAGT GGATTATGCA GACAGATGGC ACGCAGGCCT
27151 GTGTGGACAT CAGCCTCGGG CACCAGACAT AGGCAAGGCG CAAGGTGATA
27201 CAGTAGGCAG CACCATGGG GGCCAGGAGG CTCACAGAGA GGCCACACAA
27251 CCAGCCAGGA ATCCAGGACA GAGAGCTGGA ATGGAGACAG GGAAGCCAGA
27301 TACCAGGCCA GACTGGCCAG GTGCTACAGG CCTGTGGGCC AGGCCAGGCT
27351 TGGGGACTTC GTCTGGGTG TGAAGGAGAC AGGCACCCCT GAGGCCTTCC
27401 CTCTGCATCT CCAGCCCAAG CTAAGCGCAA ACTCTTAGGT TGGAGTAAGG
27451 AGTAACCCCC TGCCAAGTTT CTCCTGTCTT CAGGCTCCAC CCACCACCTA
27501 TGCTGCCTGG CCCCATGGG CACACGCTCA GGCCAGCCTT GGGAAAAGCAA
27551 CTGCACCTGC CTGTGCTATG CTGGCCCTTC TCAGCCTCAA TGCCCTCCTC
27601 CCTCCCCGAC GCACCTCGT GGCCCCCGCT GGGCCCCCTG ATGCACCCTC
27651 ATGTCTCCAT GGCAACCTGC TCAGAGTGTG GCCTTGCCCT TGGCTCCCTT
27701 CCACACCTGT GTCCAGGCA GTGCCACGGC ACTTCTCTAA ACAGAAGGAT
27751 GGGCTTCAA ACAGTCCCAG ACCTAAACA CACCTGCATT TTGGGTCCAA
27801 GTAACCTCTG ACAAGACGAG TGCCCCTACA CACCTCAGT CCTATCCACT
27851 ATGGGCAAGG AGCCTGAAGG ATCCCCCAGA ACTGGCTAAA GCCCTCAGTC
27901 TCCTCTCCA CCCTGAGCAC CTTACGCGG CAGAGTGGCC CTGGATGTCA
27951 GCTTCTTGCT CCCATGGTC TGCACCTGGA CAGGTGCTCT CAGGTGTGTG
28001 GGTGGGCAGG TGGCAGGTCC CAAGAGCCAG GTGCAAAGAA TCTAGGCCAG
28051 TGCCCACGAT TGCTGCAGTG TCTGTCCCCA GCATGGTATC TAGGGCTCCA
28101 CTTGCCTATC AGCTGTAATC GGAGGAGGCT TTCCAGGCCA GGCCTCCCCC
28151 AGGAAGGCTG CAGGCACCTG GGATCGTGCG CCCTCACATG CATTATTCTT
28201 GAGGCCCTTC TGCAGATGCC ATCAGGGCAG CAACTCTGAT GAGGTATTAG
28251 GGCACAGCAC ACAGGGCTAA GCCACCCTGT ACTGGGCCAA GCGCTACAGG
28301 CAAAAAGGAC ACCACCGACG GGCATTTTAT TCATCGCTTT TATTTTTATA

FIGURE 3, page 9 of 16

28351 TATTTTGTAG AGGGAGCCTC ACTCTGTCGC CCAGGCTGGA GTGCAGTGGC
 28401 GCGATCTTGG CTCACTGCAA CTTCTCCCTC CTGGGTTC (SEQ ID NO:3)

FEATURES:

Exon: 232-340
 Intron: 341-431
 Exon: 432-515
 Intron: 516-1110
 Exon: 1111-1205
 Intron: 1206-1424
 Exon: 1425-1547
 Intron: 1548-1981
 Exon: 1982-2065
 Intron: 2066-3015
 Exon: 3016-3058
 Intron: 3059-4102
 Exon: 4103-4177
 Intron: 4178-9088
 Exon: 9089-9126
 Intron: 9127-9303
 Exon: 9304-9375
 Intron: 9376-10898
 Exon: 10899-10943
 Intron: 10944-12713
 Exon: 12714-12762
 Intron: 12763-17130
 Exon: 17131-17133
 Intron: 17134-22868
 Exon: 22869-22944
 Intron: 22945-23137
 Exon: 23138-23154
 Intron: 23155-23475
 Exon: 23476-23705
 Stop: 23706

CHROMOSOME MAP POSITION:

Chromosome 7

ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
487	T	C	Exon	55	H	H
496	T	C	Exon	58	L	L
1662	T	C	Intron			
1785	T	A	Intron			
1889	A	T	Intron			
2416	C	T	Intron			
4698	A	G	Intron			
5424	C	T	Intron			
8722	C	A	Intron			
9982	G	A	Intron			
10951	C	T	Intron			
12603	T	C	Intron			
14583	C	T	Intron			
17290	T	C	Intron			
18188	C	T	Intron			
19911	A	G	Intron			
21328	C	A G	Intron			
21391	T	C	Intron			
22588	C	T	Intron			

22965	-	G	Intron			
23498	G	A	Exon	312	R	R
23663	T	C	Exon	367	S	S
25427	A	G	Beyond ORF (3')			
27727	C	T	Beyond ORF (3')			
27834	T	C	Beyond ORF (3')			
28336	G	A	Beyond ORF (3')			

Context:

DNA

Position

487	CACCTCTGGGTTTAAACAACATGCACCCTTGTGCCGGTCACCTCCCTGCAGCCGGAGAAC CTGCTTCTGGCCAGCAAGTGCAAAGGGGCTGCAGTGAAGCTGGCAGACTTCGGCCTAGCT ATCGAGGTGCAGGGGACCAGCAGGCATGGTTTGGTGAGTGCCAGGGGCAGGGTGTGTTG GCTGGCAGTTGGCAGGGCAGGAGGTGATGCTGACAGCCCCTTGTGGCCTCTTCCCCTCTC TCTAGGTTTCGCTGACACACAGGCTACCTGTCCCCTGAGGTCCTTCGCAAAGAGGCGTA [T, C] GGCAAGCCTGTGGACATCTGGGCATGTGGTGAGGCCTGGCCTGAGTTGGTGCGGGGCAGG GCCTCGGGTGTTTCAGGACTTCCACCTACATCCTGGAGTGTGCAGTGGCCAGCACGTCT TGCTCTCATCTGGGTTTATCTGTGTGACACCTGCCCTTGAGCTGCCCTGGCAGGGGTCTG CCCACACAGCCAAGAGCCCCCTTCCACCCAGATTAGAATTGCTCACATGAACCTGGCGC ACCCAGTGCTCGCCTGCGCTCAGCAGAGGTCTGGTCCAGAAGTGTGGTGGGTGGATGGG
496	GTTTAAACAACATGCACCCTTGTGCCGGTCACCTCCCTGCAGCCGGAGAACCTGCTTCTG GCCAGCAAGTGCAAAGGGGCTGCAGTGAAGCTGGCAGACTTCGGCCTAGCTATCGAGGTG CAGGGGACCAGCAGGCATGGTTTGGTGAGTGCCAGGGGCAGGGTGTGTTGGCTGGCAGT TGGCAGGGCAGGAGGTGATGCTGACAGCCCCTTGTGGCCTCTTCCCCTCTCTCTAGGTTT CGCTGGCACACACAGGCTACCTGTCCCCTGAGGTCCTTCGCAAAGAGGCGTATGGCAAGCC [T, C] GTGGACATCTGGGCATGTGGTGAGGCCTGGCCTGAGTTGGTGCGGGGCAGGGCCTCGGGT GTTTCAGGACTTCCACCTACATCCTGGAGTGTGCAGTGGCCAGCACGTCTTGCTCTCAT CTGGGTTTATCTGTGTGACACCTGCCCTTGAGCTGCCCTGGCAGGGGTCTGCCACACAG CCAAGAGCCCCCTTCCACCCAGATTAGAATTGCTCACATGAACCTGGCGCACCCAGTG CTCGCCTGCGCTCAGCAGAGGTCTGGTCCAGAAGTGTGGTGGGTGGATGGGAGTGGAGAA
1662	GAATTCTTGCCCCTGCCCTGAGAGGGAGCTTCAGGCCCCGCCGGGGCGCTGTTTCTTCTG CAGTTCCCGTCCCCTGAGTGGGACACCGTCACTCCTGAAGCCAAAACCTCATCAACCAG ATGCTGACCATCAACCCTGCCAAGCGCATCACAGCCCATGAGGCCCTGAAGCACCCGTGG GTCTGCGTGAGTCCGCCCTTGGTGCCCATGGTGGGGAGGGGGCTCCTGGTGGAGATGGCCT CAGACCACTCCCCTGGCAAGGACCCCAAGAGGGTCCTGTTCTGACATCCAAGAGCTCCC [T, C] TGGGTCCCCTGGGTGCTCCTTGTGGCCTCTGGCTTGGGACATACCAGCACGTTTGTGAGG CCTGGGGCTTGGAAGGCATTAGAGGGTAGAGGTGATCCCTTCTCCCAACTGCAGTCCTG TCTGTGAGGGGCAGAGTGGACGAGGCAAGGAGAGACGAGTCTTGAAGTCCCAGGCGGGT GGGGACAGACAACCTTGCCGCAATGGTGGCCGGTGGCTCTTGGCAAGTGGGGACCCAG GGTGCCACAAGCCTTGCCACCCTGGCCTCTCCCCTGTGCCTCGGGCTCGGCTGCCATATG
1785	CTGACCATCAACCCTGCCAAGCGCATCACAGCCCATGAGGCCCTGAAGCACCCGTGGGTG TGCGTGAGTCGCCCTTGGTGCCCATGGTGGGGAGGGGGCTCCTGGTGGAGATGGCCTCAG ACCAC'TCCCCTGGCAAGGACCCCAAGAGGGTCCTGTTCTTGACATCCAAGAGCTCCCCTG GGTCCCCTGGGTGCTCCTTGTGGCCTCTGGCTTGGGACATACCAGCACGTTTGTGAGGCC TGGGGCTTGGAAGGCATTAGAGGGTAGAGGTGATCCCTTCTCCCAACTGCAGTCCTGTG [T, A] GTGAGGGGCAGAGTGGACGAGGCAAGGGAGAGACGAGTCTTGAAGTCCCAGGCGGGTGGG GACAGACAACCCCTTGCCGCAATGGTGGCCGGTGGCTCTTGGCAAGTGGGGACCCAGGGT GCCACAAGCCTTGCCACCCTGGCCTCTCCCCTGTGCCTCGGGCTCGGCTGCCATATGACC ACCATTTCCCCACAGCAACGCTCCACGGTAGCATCCATGATGCACAGACAGGAGACTGT GGAGTGTCTGAAAAAGTTCAATGCCAGGAGAAAGCTCAAGGTGAGGCCCTGGCCCCTAGT
1889	GTGGAGATGGCCTCAGACCACTCCCCTGGCAAGGACCCCAAGAGGGTCCTGTTCTTGACA TCCAAGAGCTCCCTTGGGTCCCCTGGGTGCTCCTTGTGGCCTCTGGCTTGGGACATACCA GCAGGTTTGTGAGGCTTGGGGCTTGGAAGGCATTAGAGGGTAGAGGTGATCCCTTCTCTC

FIGURE 3, page 11 of 16

CAACTGCAGTCCTGTCTGTGAGGGGCAGAGTGGACGAGGCAAGGGAGAGACGAGTCTTGA
AGTCCCAGGCGGGTGGGGACAGACAACCCTTGCCGCAATGGTGGCCGGTGGCTCTTGGCA
[A, T]

GTGGGGACCCCAGGGTGCCACAAGCCTTGCCACCCTGGCCTCTCCCCTGTGCCTCGGGCT
CGGCTGCCATATGACCACCCATTTCCCCACAGCAACGCTCCACGGTAGCATCCATGATGC
ACAGACAGGAGACTGTGGAGTGTCTGAAAAGTTCAATGCCAGGAGAAAGCTCAAGGTGA
GGCCCTGGCCCCCTAGTCCCAGGCACGGCCATGCTTCTCTGTGTCCCTCTGGGCTGGAGCA
GGGGGGCCTTGGGGGGTCTGGGCAGACCTAGGGGTTACTGCTGCCCCAAGACTGACTGT

2416 TCTGGGCTGGAGCAGGGGGGCTTGGGGGGTCTGGGCAGACCTAGGGGTTACTGCTGCCC
CCAAGACTGACTGTAGCAAGTCCCAGACTGGATGCATCAGGTGAACTCAGGCCAGCTTG
GGAATGAGTCCAGAGGGGGCCTGGGCCAGGTGTGGCTCCTCCTAGTTGTCTGTGCCACCT
CCTAGCAGCCCTTGAGGAGTGTCTGAAAGCGCTCGCTGTGGGCTCCTCACCCGGGCTC
TGCAGGCAGCACTCACCCCTCTGGCAGTCACACTGTTTAGTACAAGCAAGTCCGAAGCTTC

[C, T]
GGCTCAGACAGGTTTGGAAGGAGAGCAGAGCCACACACACTGGTCTTGGGTGGGCTGGG
GGAGTTCTGGGAGGGAGGTGGGTCCAGTAGGGTATCCAACCTGCCTGCTTTGGTCAGGG
CTGGCTCCGGTGACCGCACACTGGCAGTCCCTCTACTTGTGGGTTCGGGATGGGGACTT
GTTGCCTGACTGCCCTCTGTGGTCTCTGAGCAGTTCTCCCCGGAAGCCCCAGGACTGTT
GCCCTGTCTGAGCCTGTGAGGAAAAGAGGGGCTGTGAGGGAGCTGGACCCAGAGGAGC

4698 GCTAGGTGGCCCCTGGGCTACACCAAGCCCTTCTGGTCTGGCCCCGAGGTCTGGGGGT
CCGGAGACCCCCATTAAGAATGGCCTGGGGCCCCACAGGGAGCCACTGGGCCTGCTGCTGG
GGGTCTGAATCCTGAAAGGAGAGCCTTGAGGAGCAGAGCCAGAGAGGCAGAGGCCCTTG
GGGCAGACACACCCCTGCCCTCTGGGGCCGATGGAGACGGTGGTCTGTGTGCTGAG
TCCTACACATGCATGTCTGCCCTGAGCATCCCCCAGGACAAGCCGCTCTGGAGTGGGTG

[A, G]
GGGTTTTATGCACCCCTGAGGAGACTTTCAAGGCTTCCTCTTGGGTGTTTCTGCAAAGTC
CTCCTCCCCTGGCCTCAAACCTGTGAGGGAAAAGGCCGGCACTGGCCACCTGCTCCTCT
GGGCTGTGCGGGGCAGAGCCAGAGGCCCAAGTTGGCTTCTGCCACCTGCTGGCTTGT
GACCAT

5424 CCTCCTCATGACCCACAGGGTGAGCAGCCTGGCCTTCCCAGCCAGAGAACCCTCCTTCTG
GGGAGGCCCAGGGCGTCTCGGGGAGGGCAGTCTATTCTCCTCCCATGAGCCAGTGGAC
GTGTCTAGCAGGCAGCACCCCGGGAGAGCCCTCCACGCTCTTCTCCATTTGACAGGCCTT
TCCAGAGCGCAGGCGGGAGGGGCTGTGATTAGAAAAGAGTGAGGCTAGTGGCTTCTGGG
GAGGCACTGCTGCCAGGGGACAGTGTGAGAGACAGCTGCCTCTACGCTGCCCTGTGCC

[C, T]
GGGGCTCCCGCTGCAATGCCCGCCTGTCTGCAAGTGAACGTGGGGCGACGGTGATGAGG
CCCTGCATGTGTGGCTCCACCCTGGGCGCCGAGAGCAGCTCTGTCTGGAGGGTGGTCAG
TGATGTGGACAGAGCCCAAGCATGGCTGTCTGGGTGACCAGCTAAGGGGACAAAGGCAGA
GGCAGGGCTGAGAGGACACCCATCCTGTAGGTGAGCCAGCTCAGCCATATCACACGG
CAGTGAGCATGGAGCTCAGTTCTTGCCAATGGCAGCTGAGTCTAGTACCATCCAGTCAG

8722 AAGGCCTGTGCTGGCCCCAGTCAGTGCACAGAAGCGGCCCAAGGCCAGGGCTGCTGGGC
AGCTCGGAATGAGGGCGAGCAGGGCTGCCCTTGGTGCCTGAGCCAAGGAGCCAATGGGAC
AGACCTCTGAGCCTGGGTGCCAAGTATGAGTCTGAGACAGGGTGAGCGCCTGGGCTGGG
ACAAGGCCCTCTGAGTGGGCGGCCAGCTGCAGCCACCCACCCCTACCCAGGAAGGCAG
GGCCCGGGAGGGCATGACCTCTGGGGTGTGGCTCAGCTGCCCCCACCACCACTGACAC

[C, A]
GCTAGTCCTGAGTTCCCATCAGGGAGGAAGCAGCATCCTGCCTTCCTCTAGGAAGAGCTT
GCATGTGGCCCAGAAGCCAAGGGGGCTCCCCAGCACCCACGGGCATCTCTGGGTCTGGTC
AGAGGAGAAATCTGGATGCTTGACGAGGCCCAAGGTGATGGAGAGGCTGGAGACAGGG
CTGTCTTGGGGTGATGGGATGGCCCCCACCTGCTCAGAGCCAGCCTGGGTGCTGGAAC
CACACTTGCCCTCAGGACCCTGGGCTTGCTCCTGGGGAAGAGTGGGGTCAGGCAAAGGGG

9982 CCAGGAGTGTTTCAAGGAGTCACTGAGGCAGAAAGATACCTCTCCCCACCAGGACCCACC
CTCAGCTCCTCCACCATCCTCAACAGGCCGACCCACAGACCACTCCGAAGGTCTGGCTTG
GTGGGGCTGGGCCAGGATCTGCAGGGGGAACAGCCCATAGTGGCACATTCCACGGCCCAT
GGGGAGACGGGGCCACGGTGGTGCAGTAGAGAGGTGTCTAAGCCAGTGGCAGCCAAAGGG
AGGGCTTGCCGTACCTCTGTGTCCCTCAGTGTGCTGTGTGGCTGCCTGAGAGGCAGG

[G, A]
CTTAGGGGCTCCCTGCCGGGGAGGGGAGGGGTCCCCACCATGCTCCGCTCCAAGTGGCC

FIGURE 3, page 12 of 16

CCTCAGTGCCCCCTTGCCCTGGGGGCTCCTACAGGTGAACCTATAGCAGTACTCCCAAGG
ATGTAAGATTGTGGCTGGTGGGTGCCGGCCTTCTGCTGGGGCGCTGTGCTGTGTCCCCT
CAGCTGTCTTAAGAGCTTTGGGGCTTGCTGGCCCGTAGGTCCCCATATTTGCTGGAAGCA
GGCTTGGTGTCCCCTGAGAACCCAGGCCAGGCTTCGGGAGCCAGCCCCAGACCGCCAC

10951 ACAGCAGCACCTCCGCCAGCCTGGACAGAGCTCCTGTCCATTCCATCCCTGCCGGCTGAC
CCAGGTCCTCCCCAGCTGCTCCACGCCGCTCCATCCCTGTCCCCCACTCTGCTCTGC
ACTTCTTTCTCGCAGGCTCTGGCCACCCACACCTCCTCTGTCTCCCTGTTCCCCCTCCTGG
TGGTCTCCGCTTCTCCTCTTCTCACTTTCCTCTCTTTCCTTCTCTGTGTCTTCTTC
TTCTGTAGGAGCTCAAACCACCGTCATCCATAACCCAGTGGACGGGATTAAGGTACTGC
[C, T]
CCACTTTCCTCCTCCCGCTTTCCCCAGGCAGGAGGCTCCAGGCCAGGAGAGAGGTCTGGG
GCAGCATTTGTGCCAGAGTGGAGGGCAGATGTCCCATGGCCCTGGCCGCCCTCCCCGCA
GTACGGTAGGGCCCCAGTCCGTCTTCGTGGGCAACAACAGGACAGACTGGCTCAGGCCCC
AGGCGCGCCCCTGAGGTGCTTGGCACAGTTGCGCCCGGTCCCCATGTGGCCGACACTCT
CAGACCAGGGCTCTGCGTGTCCCACCTACGGCAGGCAGTAGGGCTTCTGAGGTCTGGAG

12603 AGTCTCTCTGCCAGGCTCATCTTGCTGGGAGAAGTGGAGCCCTCATGTGTTGGGGATGCA
GGGTGGCCACAGCACTAGGGTGGCAGGGCCGGCCTCGGACTCCGTGCCAGCCTGTGCTGG
CTGCCGTGAGAATGCACCCTGGTGAGGGGCGCCCTCCAGGGACCAGCACAGAACTGGGT
GTCTTCTCCGGTCACTGCCGCATGAGGTCCACAGAGCTGGGGCCCTGCAGCCGCCAGAGG
GCATGTCCCCTGAGCCCCCTGGCCTTTAAGCCCCGTGGAAGCAGCCGAGGCAGAGATCAGC
[T, C]
TCAGAGCCTGGGCTGGTCTGTGACACAGGCCAGCCCTGTCCACCTGCCCTCAGCCACGTC
CCACCTATCCTTGCCCGCATCCTGACCCGCTGCCTCCCGTGTTCCTCAGGAGTCTTCTG
ACAGTGCCAATACCACCATAGAGGATGAAGACGCTAAAGGTACCTGCACTTGAGTCTTG
CCCCCCCAGCGGCCCTTGGCATTTGCTGGGTTGCTCTTTGAGGTGGGTGGGACTTGGGCAGG
GTCAACTCTCCTGCGACGCCCTAGTTTATGCATGTGTTGAGGGGCTCAGGGACCCTGTAGC

14583 ACATCCTGAGCTCAGTGAGGAGGGGCTCGGGAGCCCCAGAAGCCGAGGGGCCCCTGCCCT
GCCCCATCTCCGGCTCCCTTTAGCCCCCTGCCAGCCCCATGTAAGTAGCCTGGGTCTGCT
GCTGTGGGGGTGATGTTGGAGGGCTGGCAACCCCTAGAGGGGCCACTCCAGAGCCGAGG
GCAGGCTGAGCGTGGACCCTGGCTCCAGCCTCATACCCACAATCCCTCACTGGGGCTT
TCCAGGGTGGCCCCAGCCCATCGAGCCCCACCTCTTTGTGAGGAGGGCCCTGGACCACTT
[C, T]
CCTGCTCAAGGCCACTGGGCAGGATGGGAGGCCCTGGAGGCTCGGGCCTCAATTCAGTCT
TTCAGGGTGGTGAGGCCTCACTCCACCTCAGCTTGCGGGCGGGGGGCTCCCTGCTAT
TGAGGCAGGCTCTGATTAGGGCTGATCCAGGGCCCAAGGGGTCTAGAACACGGGACC
CCTCCCACTGGCTCCTCCGCTTGCCGCGCCCTCGTGTGTCTGTCTGCCTCATGTTTAC
GTCTCATCTGTTCCACCCAGCCCCAGGGATCTCTGACATCCTGAACTCTGTGAGAAGG

17290 CTGTCCCCTTGTGCCCATCCCCACATCTGCCTCTGTGCCCTCAATCTCTGGCTTGGC
TGTCTGCCCATGGTTTCTCTCCTGCGTGCCCCCGTGCTGCCTTGTGTTACGTCTCGT
CTGTTCCGCCCCAGCCCCAGGATCTCTGACATCCTGAACTCTGTGAGGAGGGGCTCAGG
GACCCAGAAGCCGAGGGCCCCCTCGCCAGTGGGGCCCCCGCCCTGCCATCTCCGACTAT
CCCTGGCCCCCTGCCACCCCATGTAAGTAGCACCTTGAGTGGCCGTGGCAGCGGTGCC
[T, C]
GGAGGGGCTCGGGGCGTGCGAGCCTGGCAGTGGTGTCTTGGGAAGGGCCATTCTTGCGGA
GGAGGGCGGGGCACAGGATCCCTCTGCTGGGTCCAGGGAATTGCTTTGAAGCACATGAA
GGTGCCACTGGGTCTCAGAAAATGGAGGTTATGGTTATGAAGTGTGTATGACATATGTGT
ATAGGAAGAGCGTCCGAAAGAGCAGGTTTGTGTGCCGACCCAGCATTCGCAACCCGAGG
TCCACAGCTTTCTCTGATGGGAGGGGAATGGGTGGCAAAGGCTCTGCGCGTGTGGCAAG

18188 ATCCCAGGGCTGTGCCACCCCCACCTGTGGGGAGACACCAGACTGGGGGTGGTGTGGAG
ATACTCTTAGAGAAGAGGCTGTGGGCCACGGGCTCGGCATGGCAGGGCAGTGGCTAGGT
AAGTACTTGAGGGACAGGTGGGGTCTGCTTGCCACCGTCCCCTCTGCAGGCTGGGCCTGG
GGGCTGTCTGCAGGCGCCAGGGCAGAAGGTTGTGGGGAGAGTGAACCCACAGGAGCAGCG
GCTCGAGGAGGGGATGCAGGCTGCAGGCTCAAAGGGGCACTGGATCCACCTGGGTGCC
[C, T]
GAGAGAGCAGGGGGCAGCCCCCTGGAGGGGTACTACCCCCAGAGCTTCTGTGGTGGCTG
AGGACCCCCAGCAGGGGTTGACTGAGGGGATCAGAGGCAAGCAGCTGAGGGGAGAGGCCA
GGTTCTTGATGCTGATAGGGTGGGGTGCCTGGGCGACCAGAACTCAAGGAGGGAGGCAT
GGGGAGGGGCGCCGTGCAGCTGGGGTGGGTGCACCGCAGAGCCTCTGGGAGTGGTCAGA

FIGURE 3, page 13 of 16

ACCCCCGACACCTGCCACTTCTACAGCAGCTCATCTGATTTTAAGGGGCTTGCTGCCCTT
 19911 AGCACGGTTACCACTCTTGATTGGAACCTGACCATGCATCTCCTCTTCTGTTACTTCA
 CGCTTTCTCTTCCCATCAACTCCCATTTTAATTACAATTTGTTTAAAAGCACTGCATATT
 ACTTCATTAAACAGAAGATTAGTTTCACTTACCATTAGTGTAAGGTGACTATAGAACCAA
 AGCAGACTGGAAACCAATGACATAATGTATTCTTCTTCCATTCCAGCTGCCTGCTGC
 TGTGCGCCTGAGAACCCCTGTGGAGTGGGAGGGGAGCTGTCTCTGTACATTAGAAAGGG
 [A, G]
 GGTTAACATAAGTGACAGGAGGTGTTTGGGACATGTGGACACCAGACTTCTCTCTTGATGC
 AAGGAGGGCAGAGCCAGGCAGCCTAGTGGGGGCTGGCTTGGGGGCTGCTGGAAGGACTGG
 CTACAGGTGGAAGAGAGGTGACACCTGAAGCTTGGGGCCACCTCCAGGAAAGGACAGGTG
 AAAGTGGAGGCATGAGGCAGGGGACAGGCAGGTGCCAGGCAGAGGGTGGAGAGGAGGCAG
 GAACATAGCAGCTGGGGCGGGGCGGGCCCTCAAGTGTATATGCTACTTTCTTGGGGCC
 21328 GCTGGGCACAGTGGCTCATACCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGAT
 CACTTGAGGTTAGGAGTTTGAGACCAGCCTGGCCAATATGGTGAAACCTCATCTCCACTA
 AAAATATACACACACAAAAATTAGCTGGGTGTGGTGGTGTGCACCTGTAGTTCCAGCTAC
 TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCTGGGAGTGCAGAGACTGCAGTGAGCCGA
 GATCATGTCACTGCACTCCAGCCCGGTGACAGAGTGAGACTCCATCTAAAAA
 [C, A, G]
 AATTCCTCTCTCTGGGAATTTAGACCACAGACAGGTTGCATGTATGTGGCCGTTGGAGGC
 AGCACTCACAGCAAAGAGTGGAAACGTCAACACAGGGCCTGCCTTCTGGTGAAAATGGTG
 TCCTGCAGGGCGGGCAGCTGTTTGAGGGCAGGTGTCCAGGTGCGGCCCTGCAGCAGCCTG
 AGGGTCACAGAGCGCAGTGCTGGGAGTGCAGAGACTTCCCCACAGGGAGAGTTCCCAGG
 AACCTGCTTCCGGTGCACCTTCTGGGGGTTTGTAGTTTTTCCACGGACGAATTACTTTGAG
 21391 TTGAGGTTAGGAGTTTGAGACCAGCCTGGCCAATATGGTGAAACCTCATCTCCACTAAAA
 ATATACACACACAAAAATTAGCTGGGTGTGGTGGTGTGCACCTGTAGTTCCAGCTACTCG
 GGAGGCTGAGGCGAGGAGAAATCGCTTGAACCTGGGAGTGCAGAGACTGCAGTGAGCCGAGAT
 CATGTCACTGCACTCCAGCCCGGTGACAGAGTGAGACTCCATCTAAAAA
 [T, C]
 ACTCACAGCAAAGAGTGGAAACGTCAACACAGGGCCTGCCTTCTGGTGAAAATGGTGTCC
 TGCAGGGCGGGCAGCTGTTTGAGGGCAGGTGTCCAGGTGCGGCCCTGCAGCAGCCTGAGG
 GTCACAGAGCGCAGTGCTGGGAGTGCAGAGACTTCCCCACAGGAGAGTTCCAGGAAC
 CTGCTTCCGGTGCACCTTCTGGGGGTTTGTAGTTTTTCCACGGACGAATTACTTTGAGAAA
 CCACTGTTACTCGTGTGTATAGGTGAGCGTGCCTGTGCATGTGTGTCTGTGTGTGAGTG
 22588 GCTGCTTCCTCCTCCCCGGCCTCCGGGTGGCCTTGCTGACGGCTCCTTCTCTGAGGCAGG
 TCTCTGCCTTCTCGCTGGTGCCTGCACCTAGTACCCCTCACCAGAGCTGCTGGGTGA
 AGGAAGCACTAAGAACCCAGGCTCGGGAGGAGAGTGGGGCCGGGAAGCTGCAGGGAAGC
 GCAGGGCCAGGCCCTGGTGGGGCCAGGGGCTGGCTCACGGGAGGGCAGGAGGGAGACTGTG
 GCGGACAGCACGTGGGGCCAGGAGGTGACCTCCAAGTGGATTGTGGGTGGGTTTTTTGTG
 [C, T]
 TCTTTCTGCATTTTCCAGGCATTTTGTAAATGTGGATAGAATATTTCTGTTCTTCAAAAAT
 ACTTTAGTTAAGAAAAATAAGATGGAAGCTGTTGCACTTGAAAATGAGGAAGCCACTGGT
 GATGCAGGGGGGCGCGCGGAGAGGACCTTCTGCAAAATAGCGGCAGGAACACGGCATGG
 ATGCAGCTCGCGCTCCCCAGGCCCTCCCCTGGGCTGTGTGGAGGGGTCCGGGGGGAATG
 GGCCAGCGCCAGTGGTCACCTGGCCATGTCTCCCCACAGCCCGGAAGCAGGAGATCATT
 22965 ATAAGATGGAAGCTGTTGCACCTGAAAATGAGGAAGCCACTGGTGATGCAGGGGGGGCGG
 CGGAGAGGACCTTCTTGCAATAGCGGCAGGAACACGGCATGGATGCAGCTCGCGCTCC
 CCCAGGCCCTCCCCTGGGCTGTGTGGAGGGGTCCGGGGGGAATGGGCCAGCGCCAGTGG
 TCACCTGGCCATGTCTCCCCACAGCCCGGAAGCAGGAGATCATTAAAGACCACGGAGCAGC
 TCATCGAGGCCGTCAACAACGGTGACTTTGAGGCCTACGCGTGTAGTCCCTGGGGCTGGGG
 [-, G]
 GGGGCTGTGCAGGACAAGGATGTGGGACCCCTGGGGGGGCTGCTCAGAGTCAGGGGTCC
 ACGGGGCCCCCTCCTCACTTGGATTGTGGCCCCCAGGAAAATCTGTGACCCAGGGCTGACCT
 CGTTTGAGCCTGAAGCACTGGGCAACCTGGTTGAAGGGATGGAATCCACAGATTCTACT
 TCGAGAACCCTGAGTGAGGAAGCCCGGGTGGGCATGAGGGGGCGGTGCCCCCAGGAGAGC
 CTCTCGGCCCTCCCAGGGACAGCATGGTGGCTGCCTATGGAAGCCCTGTCCCCTCTGTG
 23498 CCCGCCAGAGGCCATACCCAGCCCCAGAATCCCACCTTGGAGGGGCCCATGCTGCTCC

FIGURE 3, page 14 of 16

CAGGAGAGCCGAGCCTCCCCAATAAGGGGAGTTGAGAGAGGGAAAGGATTAGGCTGGTGG
GGTGAAGACGGGCACACAGGGCAGTCATGGTAACCCGAGACCCCCGCCCCGCTGCTGTC
CACAGTGCTGGCCAAGAACAGCAAGCC

[G, A]

ATCCACACGACCATCCTGAACCCACACGTGCACGTCAATTGGAGAGGATGCCGCTGCATC
GCTTACATCCGGCTCACGCAGTACATTGACGGGCAGGGCCGGCCCCGCACCAGCCAGTCT
GAGGAGACCCGCGTGTGGCACCGCCGACGGCAAGTGGCAGAACGTGCACTTCCACTGC
TCGGGCGCGCCTGTGGCCCCGCTGCAG

23663

GCCTCCCCAATAAGGGGAGTTGAGAGAGGGAAAGGATTAGGCTGGTGGGTGGAAGACGG
GCACCAGGGCAGTCATGGTAACCCGAGACCCCCGCCCCGCTGCTGTCCACAGTGCTGGC
CAAGAACAGCAAGCCGATCCACACGACCATCCTGAACCCACACGTGCACGTCAATTGGAGA
GGATGCCGCTGCTACGCTTACATCCGGCTCACGCAGTACATTGACGGGCAGGGCCGGCC
CCGCACACGACAGTCTGAGGAGACCCGCGTGTGGCACCGCCGACGGCAAGTGGCAGAA

[T, C]

GTGCACCTTCCACTGCTCGGGCGCGCCTGTGGCCCCGCTGCAGTGAAGGTGAGTGTCTGT
GCTAAGTGACAGCTGGGGCAGAGGGGTGGCGGTGGTGTGAGTGGCTGCAGCCTGGGGAGG
CGATGGGGAGCGGTGGGGCCTGTGGCAGAGCCCATGCCTGGGAAGTCCCTGAGCTTTCCT
GGTGAGGCCACAGGAATGATGTCAAATTAGGGACCACGGCAGGCTGGGTGTGGCAGGCCT
CCCCAGAGGACTGGGGAGCTGGTGAGGGCCTGAGCAGTCCACACTGGCCAGAGCTGGGTG

25427

TGTGGCAAGAGGACTCTGCCTGGGCTGGCCCCCTCCTGTGTGAGGTGTCTGTCCCTTCT
CTGCTGGCCAGCAGCAGATGCACTGGCAGCTCCCAACCTGTTTCCGCCCCCTCGGCCCTC
CCCCAGCCTGTTCCGCTTCTCTGCAGCCCGCAAGGGGGAGCAGACTTTTGACAAAGGACT
GCGGGCCTCGCTCAAGTCCCTGAGCCCCCAGCTGAAGCTGGGAGGGGAGGCCAGGCTTTG
TGTCTGGGCATATTCGTCTGCTGATGGGGTTTGGGGAAGCCTGGGGCTTGGGGTTTGGTC

[A, G]

GGTGGTGACAGCTAGTGGCAGAGCGGGATCAGAGGTGGTGGCTGCCAGCTTCTGGGCTGA
GACAAGGGTCTGTGCAGGGGTTTACTGAAGTGGGAGTGCCTTTGGAATCTGGGCCGGGAG
CAGAAGGGAGCAAAAGCTACAGTGGGAGCCAGCCTAGGGCACATGGGAGGCGTGAGGGCA
GTGCTGCCCGTGCAGTGTGAGGTGTGCCAGTGCCTTGGCGGGCTGCAGTGCCTGTGAGGG
CACCTTCTAGGTGGGCCAGGGATGCAGCTATGGAGATAAGGCGGGCTGGGGACAGAAACA

27727

GCAAACCTCTTAGGTTGGAGTAAGGAGTAACCCCTGCCAAGTTTCTCCTGTCTCAGGCT
CCACCCACACCTATGCTGCCTGGCCCCATGGGGCACACGCTCAGGCCCAGCCTGGGAAA
GCAACTGCACCTGCCTGTGCTATGCTGGCCCTTCTCAGCCTCAATGCCCTCCTCCCTCCC
CGACGCACCCCTCGTGGCCCCCGCTGGGCCCCCTGATGCACCCCTCATGTCTCCATGGCAAC
CTGCTCAGAGTGTGGCCCTGCCCTTGGCTCCCTCCACACCTGTGTCCAGGCAGTGCCA

[C, T]

GGCACCTTTCCTAACAGAAGGATGGGCTTCAAAACAGTCCCAGACACTAAACACACCTGC
ATTTTGGGTCCAAGTAACCTTCTGACAAGACGAGTGCCCTACACACCCTCAGTCCCTATCC
ACTATGGGCAAGGAGCCTGAAGGATCCCCCAGAACTGGCTAAAGCCCTCAGTCTCCTCCT
CCACCCCTGAGCACCTTCACGCGGCAGAGTGGCCCTGGATGTGAGCTTCTTGCTCCCCATG
GTCTGCACCTGGACAGGTGCTCTCAGGTGTGTGGGTGGGCAGGTGGCAGGTCCCAAGAGC

27834

CCAGCCTGGGAAAGCAACTGCACCTGCCTGTGCTATGCTGGCCCTTCTCAGCCTCAATGC
CCTCCTCCCTCCCCGACGCACCCCTCGTGGCCCCCGCTGGGCCCCCTGATGCACCCCTCATG
TCTCCATGGCAACCTGCTCAGAGTGTGGCCCTGCCCTTGGCTCCCCCTCCACACCTGTGTC
CCAGGCAGTGCCACGGCACTTTCCTAAACAGAAGGATGGGCTTCAAAACAGTCCCAGACA
CTAAACACACCTGCATTTTGGGTCCAAGTAACCTTCTGACAAGACGAGTGCCCCCTACACAC

[T, C]

CTCAGTCTCTATCCACTATGGGCAAGGAGCCTGAAGGATCCCCCAGAACTGGCTAAAGCCC
TCAGTCTCCTCCTCCACCCTGAGCACCTTCACGCGGCAGAGTGGCCCTGGATGTGAGCTT
CTTGCTCCCCATGGTCTGCACCTGGACAGGTGCTCTCAGGTGTGTGGGTGGGCAGGTGGC
AGGTCCCAGAGCCAGGTGCAAAGAATCTAGGCCAGTGGCCACGAGTGTGCTGAGTGTCTG
TCCCCAGCATGGTATCTAGGGCTCCACTTGCCTATCAGCTGTAATCGGAGGAGGCTTTCC

AAGAATCTAGGCCAGTGCCTCCAGAGTGTCTGTCAGTGTCTGTCCCCAGCATGGTATCTAGGG
 CTCCACTTGCCTATCAGCTGTAATCGGAGGAGGCTTTCCAGGCCAGGCCCTCCCCAGGAA
 GGCTGCAGGCACCTGCGGATCGTGCGCCCTCACATGCATTATTCCTGAGGCCCTTCTGCAG
 ATGCCATCAGGGCAGCAACTCTGATGAGGTATTAGGGCACAGCACACAGGGCTAAGCCAC
 CCTGTACTGGGCCAAGCGCTACAGGCAAAAAGGACACCACCGACGGGCATTTCATTCATC
 [G,A]
 CTTTATTATTTATATATTTTGGAGAGGGAGCCTCACTGTGCGCCAGGCTGGAGTGCAG
 TGGCGCGATCTTGGCTCACTGCAACTTCTCCCTCCTGGGGTTC

[illegible]